Università degli Studi di Genova
Facoltà di Medicina e Chirurgia

UOC Clinica Pediatrica e Reumatologia
IRCCS Istituto Giannina Gaslini

Tesi di Dottorato in Scienze Pediatriche
XXXII Ciclo – Curriculum Reumatologia Pediatrica

Development of a prognostic model for Macrophage Activation Syndrome in Systemic Juvenile Idiopathic Arthritis

Relatore:

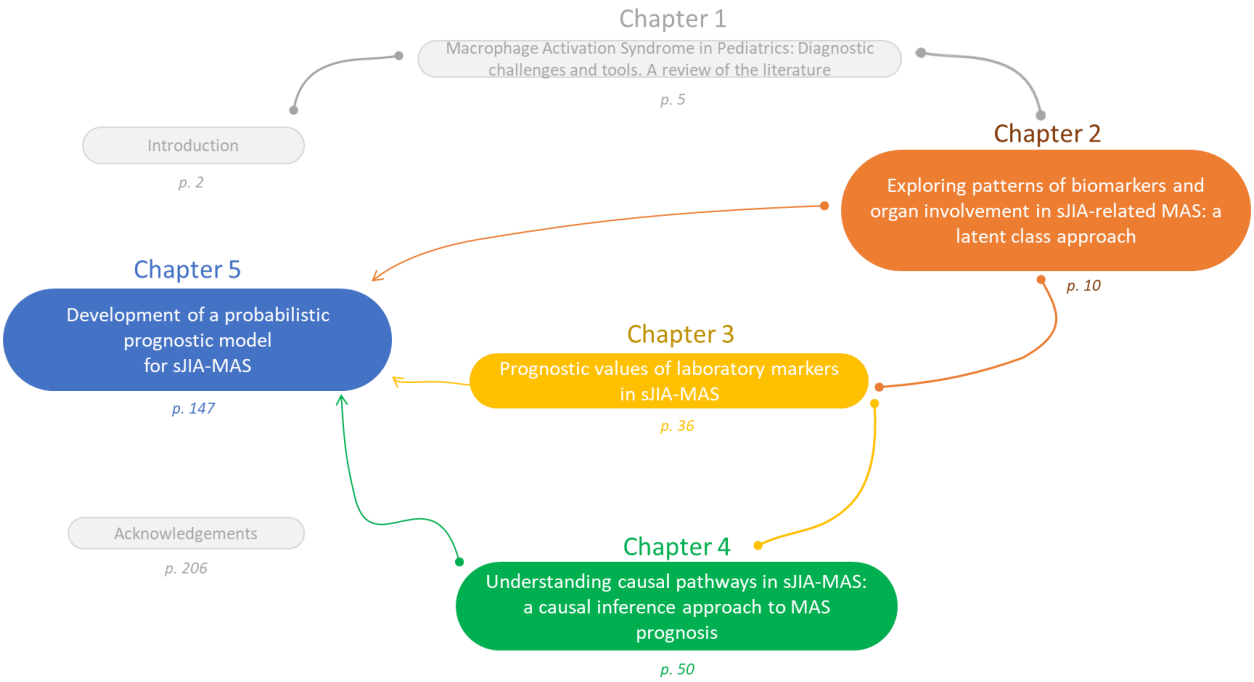
Chiar.mo Prof. Angelo Ravelli

Candidato:

Dott.ssa Alessandra Alongi

Anno Accademico 2019/2020

Index



Introduction

> He maintained, among other things, that unanticipated disasters are never the result or, say, the effect of a single motive, of a one and only cause, but resemble a vortex, a point of cyclonic depression in the world's consciousness: a spot towards which a multiplicity of conspiring factors have converged.
>
> *Carlo Emilio Gadda. Quer pasticciaccio brutto de via Merulana. Garzanti, 1970*

Macrophage activation syndrome (MAS) is a severe and potentially life-threatening complication of many systemic diseases that shows a particular association with Systemic Juvenile Idiopathic Arthritis (sJIA). MAS belongs to the spectrum of Hemophagocytic Lymphohistiocytoses (HLH), a group of both genetic and acquired conditions characterized by an hyperinflammatory status driven by the uncontrolled secretion of pro-inflammatory cytokines due to the proliferation of activated lymphocytes and histiocytes[1].

This systemic, widespread inflammation can result in a wide range of clinical manifestations and laboratory abnormalities, including – but not limited to - hyperferritinemia, pancytopenia, liver and coagulation dysfunction and central nervous system involvement. In the most severe patients, renal, pulmonary and cardiac involvement may occur, leading to multiple organ failure and, eventually, to a fatal outcome[2]. The clinical course of these patients can in some cases be rapid or fulminant and is generally hardly predictable. Moreover, the grade of organ involvement and clinical and laboratory manifestations are highly heterogenous[3]. Despite the recognized

[1] Minoia, Francesca, et al. "Criteria for Cytokine Storm Syndromes." Cytokine Storm Syndrome. Springer, Cham, 2019. 61-79.
[2] Ravelli, Angelo, et al. "Macrophage activation syndrome." Hematology/Oncology Clinics 29.5 (2015): 927-941.
[3] Minoia, Francesca, et al. "Dissecting the heterogeneity of macrophage activation syndrome complicating systemic juvenile idiopathic arthritis." *The Journal of rheumatology* 42.6 (2015): 994-1001.

severity of the condition, factors responsible of the poor outcome observed among a subgroup of MAS patients have not been extensively investigated. In fact, the majority of prognostic studies on the HLH/MAS spectrum concentrates on the more severe familiar HLH forms, while few evidences on the prognosis of sJIA-related MAS are avaible. Yet, the possibility to predict the clinical course at disease diagnosis could have obvious implications for the prompt initiation of more aggressive treatments and strict monitoring of high-risk patients.

While recent advances have been made on the identification of molecular and cytokine-related biomarker for the diagnosis and monitoring of disease activity[4], most of these tests are not easily obtained in clinical practice and not avaible in most centers.

In this project, we aimed to develop a prognostic model for sJIA-MAS based on routinely avaible laboratory parameters collected at MAS onset. For this task, we conducted a re-analysis of the dataset gathered in a multinational, multicenter project coordinated by the PRINTO International Centre at Gaslini Institute, in Genoa. The study involved a large number of pediatric rheumatologists and haemato-oncologists worldwide, with the original aim to develop a set of classification criteria for sJIA-MAS[5]. To account for the complexity and heterogeneity of MAS patients, we proceeded through separate analytical steps that were finally integrated in a probabilistic risk model. We focused on the prediction of severe outcome, defined in the original study as a composite outcome including mortality and admission in Intensive Care Unit, death and the development of central nervous system (CNS) dysfunction. The objectives and results of these works, which are reported in the various chapters of this thesis, are recapitulated below.

In ***Chapter 1***, a literature review of MAS is presented, covering the recent efforts in diagnostic and classification approaches. This work has been accepted for publication in the journal

[4] Mizuta, Mao, et al. "Clinical significance of serum CXCL9 levels as a biomarker for systemic juvenile idiopathic arthritis associated macrophage activation syndrome." Cytokine 119 (2019): 182-187.

[5] Ravelli, Angelo, et al. "2016 Classification criteria for macrophage activation syndrome complicating systemic juvenile idiopathic arthritis: a European league against rheumatism/American College of Rheumatology/Paediatric Rheumatology International Trials Organisation Collaborative Initiative." Arthritis & Rheumatology 68.3 (2016): 566-576.

"Pediatric Allergy and Immunology" with the titled "Macrophage Activation Syndrome in Pediatrics".

In ***Chapter 2****,* we report an explorative subgroup analysis conducted by the means of Latent Class Modelling, a method allowing for the identification of homogenous subsets of patients, based on selected predictors avaible at diagnosis, among heterogenous populations. The associations with baseline features and outcomes are investigated

In ***Chapter 3*** we assess the prognostic values of routine laboratory parameters at MAS onset to identify the most predictive factors by regression modelling. A theoretic, knowledge-based causal model for each outcome, formalized as a direct acyclic graph, is employed to addressing possible bias.

In ***Chapter 4*** we further investigate the effects of specific predictors on outcomes to identify the causal determinants of clinical course in MAS. For this aim, we employ a causal inference approach which allow to estimate the unbiased impact of each risk factor and to identify the aspects that need to be targeted to improve the outcome.

In ***Chapter 5*** we integrate the previous results in the final prognostic model and related score, developed within a probabilistic framework. By means of Bayesian Network methodology and constructs from information theory, we identify the most impacting and informative risk factors, and present the implementation of a digital tool which provides personalized predictions and can assist physicians in clinical decision-making.

In ***Chapter 6*** we draw the conclusions from the presented work and outline some possible directions of future research.

Chapter 1 – Macrophage Activation Syndrome in Pediatrics: Diagnostic challenges and tools. A review of the literature

Macrophage Activation Syndrome (MAS) is a potentially life-threatening hyperinflammatory condition belonging to the spectrum of hemophagocytic lymphohistiocytosis (HLH) that can complicate many immunological and rheumatic disorders. HLH describes a group of histiocytic disorders driven by the uncontrolled activation of T cells and macrophages exhibiting haemophagocytic activity. The current classification of histiocytic disorders includes MAS among the forms of HLH occurring in the context of pre-existing disorders, such as infectious, rheumatic and oncological diseases – namely secondary or acquired HLH - as opposed to primary or familiar forms in which known genetic defects can be identified. Among rheumatic disease, the highest predisposition to developing MAS is observed in patients with systemic juvenile idiopathic arthritis (sJIA), 10–30% of which will develop clinical or subclinical MAS at some point of the disease course; less frequently, MAS can also occur in Juvenile Systemic Lupus Erythematosus, with a reported prevalence of 0.9 to 4.6 %, Kawasaki Disease and Juvenile Dermatomyositis[6].

[6] Ravelli, Angelo, et al. "Macrophage activation syndrome." Hematology/Oncology Clinics 29.5 (2015): 927-941.

Current pathogenetic views suggest that primary and secondary HLH may represent a spectrum in which different mechanisms at multiple levels converge in the self-perpetuating activation of T cells and macrophages, sustained by the hyperproduction of IFNγ. In primary HLH, this scenario is maintained by the alteration of immunoregulatory feedback mechanisms due to genetic defects involving the perforin-mediated cytotoxicity pathway. In MAS, multiple mechanisms can contribute to macrophage and T-cells hyperactivation, including infectious triggers, hypomorphic genetic mutations, and the specific cytokine milieu characterizing some rheumatologic disorders. Specifically, the unbalanced production of biologically active IL-18 seems to play a major role in predisposing patients with sJIA, Kawasaki Disease and some autoinflammatory disorders to MAS. In all these forms, the common endstage of these different pathways will be the overwhelming release of pro-inflammatory mediators such as IL-1, IL-6, TNFa - sometimes described as "cytokine storm" - resulting in similar clinical phenotypes[7].

Clinically, MAS is characterized by the abrupt onset of unremitting high fever, pancytopenia, hyperferritinemia, liver and coagulation dysfunction, eventually progressing, if untreated, in multi-organ failure and possible fatal outcome. Encephalopathy is reported in around one-third of cases. Bone marrow hemophagocytosis, while supporting the diagnosis, shows low sensitivity (approximately 0.60); thus, its absence does not rule out MAS/HLH1.

Different clinical pictures - including sepsis, viral disease (most frequently EBV), infection from intracellular pathogens (Leishmania spp.) - can both mimic and MAS. In rheumatic patients, the substantial overlapping between clinical and laboratory features of MAS and the predisposing

[7] Bracaglia, Claudia, Giusi Prencipe, and Fabrizio De Benedetti. "Macrophage activation syndrome: different mechanisms leading to a one clinical syndrome." Pediatric Rheumatology 15.1 (2017): 5.

condition complicates the early recognition of MAS from flares or infections. Moreover, while clinically similar to MAS, genetic forms of HLH may warrant specific recognition as they are generally more severe, thus requiring more aggressive therapeutic approaches.

To address these challenges, in the last years a great effort has been made to develop diagnostic tools for identifying MAS in specific rheumatic diseases, focusing mostly on sJIA. Generic diagnostic guideline for HLH, such as the HLH-2004 diagnostic criteria, have demonstrated low sensitivity for MAS in the context of rheumatic conditions, particularly in sJIA. Therefore, an international collaboration was established to develop specific diagnostic criteria for MAS complicating sJIA. Basing on both expert consensus and data from over 1000 sJIA patients, of which 362 diagnosed with MAS, preliminary diagnostic guidelines aimed at differentiating MAS from sJIA flare and systemic infections were proposed in 2005[8]. A multistep expert consensus process finally led to the development and validation of the 2016 Classification Criteria; according to this new set, the diagnosis of MAS in a febrile patient with known or suspect sJIA requires serum ferritin level > 684 ng/ml and at least two among the following criteria: platelet count ≤ 181 × 109/l, aspartate aminotransferase > 48 units/l, triglyceride concentration > 156 mg/dl, fibrinogen ≤ 360 mg/dl[9]. Although altogether sensitive (0.73) and specific (0.99), these criteria may be of limited use for the diagnosis of early or subclinical form of MAS in clinical settings[10]. In these cases, it has been suggested that the temporal dynamics of laboratory biomarkers may be

[8] Ravelli, Angelo, et al. "Preliminary diagnostic guidelines for macrophage activation syndrome complicating systemic juvenile idiopathic arthritis." The Journal of pediatrics 146.5 (2005): 598-604.
[9] Ravelli, Angelo, et al. "2016 Classification criteria for macrophage activation syndrome complicating systemic juvenile idiopathic arthritis: a European league against rheumatism/American College of Rheumatology/Paediatric Rheumatology International Trials Organisation Collaborative Initiative." Arthritis & Rheumatology 68.3 (2016): 566-576.
[10] Shimizu, Masaki, et al. "Validation of classification criteria of macrophage activation syndrome in Japanese patients with systemic juvenile idiopathic arthritis." Arthritis care & research 70.9 (2018): 1412-1415.

more informative than the achievement of an absolute threshold. Specifically, an expert panel within the same collaborative project identified platelet count, ferritin, and AST as the parameters whose change over time is more relevant for the diagnosis of sJIA-related MAS[11]. Moreover, as shown in a recent systematic review[12], the 2016 criteria may underperform for patients who develop MAS under IL-1 blocking treatment, which are less likely to be febrile, or those under anti-IL-6 medications, which show tendentially lower ferritin and platelets levels. As MAS occurs at the onset of SJIA in about 20% of cases[13] and molecular testing is not readily available in some clinical contexts, the prompt differentiation from primary HLH is not always straightforward. A specific index - the MH score- developed by comparing clinical and laboratory features of 362 patients with sJIA-related MAS and 258 patients with primary HLH, has been proposed to help physicians to discriminate the two conditions. The score includes age at onset ≤1.6y, splenomegaly, neutrophil count ≤1.4 x109/liter, platelet count ≤78 x109/liter, fibrinogen ≤131 mg/dl and hemoglobin ≤8.3 g/dl as most suggestive criteria for primary HLH[14]. Future prospective validation studies are warranted to evaluate the utility of these diagnostic tools in various settings.

In summary, MAS is a severe complication occurring in rheumatic diseases, especially sJIA. Recent international efforts have provided several consensus tools for the differentiation of MAS from

[11] Ravelli, Angelo, et al. "Expert consensus on dynamics of laboratory tests for diagnosis of macrophage activation syndrome complicating systemic juvenile idiopathic arthritis." RMD open 2.1 (2016): e000161.
[12] Schulert, Grant S., et al. "Effect of biologic therapy on clinical and laboratory features of macrophage activation syndrome associated with systemic juvenile idiopathic arthritis." Arthritis care & research 70.3 (2018): 409-419.
[13] Minoia, Francesca, et al. "Clinical features, treatment, and outcome of macrophage activation syndrome complicating systemic juvenile idiopathic arthritis: a multinational, multicenter study of 362 patients." Arthritis & Rheumatology 66.11 (2014): 3160-3169.
[14] Minoia, Francesca, et al. "Development and initial validation of the macrophage activation syndrome/primary hemophagocytic lymphohistiocytosis score, a diagnostic tool that differentiates primary hemophagocytic lymphohistiocytosis from macrophage activation syndrome." The Journal of pediatrics 189 (2017): 72-78.

confounding diseases and primary HLH, thus assisting diagnosis and enabling future research. Together with novel advances in pathogenesis and therapeutic options, these advances are expected to improve the outcome of this condition in the future.

References

- Bracaglia, Claudia, Giusi Prencipe, and Fabrizio De Benedetti. "Macrophage activation syndrome: different mechanisms leading to a one clinical syndrome." Pediatric Rheumatology 15.1 (2017): 5.
- Minoia, Francesca, et al. "Clinical features, treatment, and outcome of macrophage activation syndrome complicating systemic juvenile idiopathic arthritis: a multinational, multicenter study of 362 patients." Arthritis & Rheumatology 66.11 (2014): 3160-3169.
- Minoia, Francesca, et al. "Development and initial validation of the macrophage activation syndrome/primary hemophagocytic lymphohistiocytosis score, a diagnostic tool that differentiates primary hemophagocytic lymphohistiocytosis from macrophage activation syndrome." The Journal of pediatrics 189 (2017): 72-78.
- Ravelli, Angelo, et al. "2016 Classification criteria for macrophage activation syndrome complicating systemic juvenile idiopathic arthritis: a European league against rheumatism/American College of Rheumatology/Paediatric Rheumatology International Trials Organisation Collaborative Initiative." Arthritis & Rheumatology 68.3 (2016): 566-576.
- Ravelli, Angelo, et al. "Expert consensus on dynamics of laboratory tests for diagnosis of macrophage activation syndrome complicating systemic juvenile idiopathic arthritis." RMD open 2.1 (2016): e000161.
- Ravelli, Angelo, et al. "Macrophage activation syndrome." Hematology/Oncology Clinics 29.5 (2015): 927-941.
- Ravelli, Angelo, et al. "Preliminary diagnostic guidelines for macrophage activation syndrome complicating systemic juvenile idiopathic arthritis." The Journal of pediatrics 146.5 (2005): 598-604.
- Schulert, Grant S., et al. "Effect of biologic therapy on clinical and laboratory features of macrophage activation syndrome associated with systemic juvenile idiopathic arthritis." Arthritis care & research 70.3 (2018): 409-419.
- Shimizu, Masaki, et al. "Validation of classification criteria of macrophage activation syndrome in Japanese patients with systemic juvenile idiopathic arthritis." Arthritis care & research 70.9 (2018): 1412-1415.

Chapter 2 - Exploring patterns of biomarkers and organ involvement in sJIA-related MAS: a latent class approach

1. Introduction

Macrophage activation syndrome (MAS), also known as secondary hemophagocytic lymphohistiocytosis (HLH), describes a severe hyperinflammatory condition that can complicate various systemic illness (e.g., autoimmune disease, malignancy, infection), particularly systemic juvenile idiopathic arthritis (SJIA)[15].

The underlying immunopathological mechanisms, which involve an uncontrolled release of immune mediators by macrophages and T cells known as a "cytokine storm", lead to a wide spectrum of clinical and laboratory features including sustained fever, hyperferritinemia, pancytopenia, hepatobiliary dysfunction, consumptive coagulopathy resembling disseminated

[15] Ravelli, Angelo, et al. "Macrophage activation syndrome." Hematology/Oncology Clinics 29.5 (2015): 927-941.

intravascular coagulation, and central nervous system (CNS) involvement. If undertreated, MAS can evolve into multi-organ failure and ultimately death.

Despite the recent advances in translational research, which identified IL-18 and IFNγ as driving mediators of the syndrome[16], the pathophysiological mechanisms leading to organ dysfunction in MAS patients are not fully elucidated. The development and evolution of this condition is known to be influenced by many factors, including the activity status of the underlying disease, received treatments, infection triggers, and genetic alterations[17]. As a result, presentation and severity differ from patient to patient, and the clinical course of can be hardly predictable, with a rapid and sometime fulminant progression in the most severe cases[18]. However, easily avaible predictors able to distinguish subject who will develop severe organ involvements have not been identified yet. To improve the outcome prediction and identify high-risk subgroup of patients, the heterogeneity of MAS needs to be addressed. To explore comprehensively the characteristic of these patients at presentation and detect potential laboratory signature that can identify subjects with poor outcomes, we analyzed the largest multinational SJIA-related MAS cohort collected so far by means of latent class modelling. This a model-based exploratory analysis allows to identify multiple unobserved clusters in complex populations by estimating categorical latent variables defining homogeneous subgroups. This approach is widely used in different field

[16] Crayne, Courtney B., et al. "The Immunology of Macrophage Activation Syndrome." Frontiers in immunology 10 (2019).
[17] Bracaglia, Claudia, Giusi Prencipe, and Fabrizio De Benedetti. "Macrophage activation syndrome: different mechanisms leading to a one clinical syndrome." Pediatric Rheumatology 15.1 (2017): 5.
[18] Minoia, Francesca, et al. "Criteria for Cytokine Storm Syndromes." Cytokine Storm Syndrome. Springer, Cham, 2019. 61-79.

to characterize and stratify patients into risk groups[19]. Assessing the predictive value of laboratory patters at disease onset on the outcomes may improve the recognition of patients who are likely to develop a life-threatening disease and thus identify those who would benefit most from escalation of treatment.

2. Materials and Methods

2.1 Study population

A multinational cohort of 362 patients diagnosed with sJIA-MAS was analyzed. Data were gathered through a web-based survey and included laboratory parameters collected at MAS diagnosis, distal outcomes such as mortality and "severe course", defined as a composite outcome given by the occurrence of death or ICU admission. Demographical features Organ involvement develop through the clinical course and administered treatments were detailed for each patient. Detailed characteristics of this cohort have been reported previously[20].

2.2 Statistical Methods

Latent Class (LCA) modeling is a probabilistic technique that uses mixture modelling to identify unobserved, homogeneous subgroups of patients (classes or clusters) in an heterogenous population, based on the conditional response probabilities of selected observable predictors[21].

[19] Campbell, Christina A., et al. "Assessing Intervention Needs of Juvenile Probationers: An Application of Latent Profile Analysis to a Risk–Need–Responsivity Assessment Model." Criminal Justice and Behavior 46.1 (2019): 82-100.
[20] Minoia, Francesca, et al. "Clinical features, treatment, and outcome of macrophage activation syndrome complicating systemic juvenile idiopathic arthritis: a multinational, multicenter study of 362 patients." Arthritis & Rheumatology 66.11 (2014): 3160-3169.
[21] Muthén, B. (2004). Latent variable analysis. In D. Kaplan (Ed.), The Sage handbook of quantitative methodology for the social sciences (pp. 345–368). Thousand Oaks, CA: Sage Publications.

Model fit statistics are used to identify the most appropriate number of latent classes[22]. Posterior probabilities indicating the likelihood of membership into each class are estimated. Once identified, the classes can be investigated in relation to their association with other variables (not included in the model definition) or distal outcomes.

In this analysis, data were preprocessed by removing highly correlated variables (< Spearmen's r = 0.5). Hemoglobin, whole blood cells, platelets, ERS, CRP, AST, bilirubin, fibrinogen, d-dimer, ferritin, BUN and creatinine values avaible at MAS diagnosis were included as class-defining variables in the LCA model. Missing data were imputed using multiple imputation procedures.

[22] Berlin, Kristoffer S., Natalie A. Williams, and Gilbert R. Parra. "An introduction to latent variable mixture modeling (part 1): Overview and cross-sectional latent class and latent profile analyses." Journal of Pediatric Psychology 39.2 (2014): 174-187.

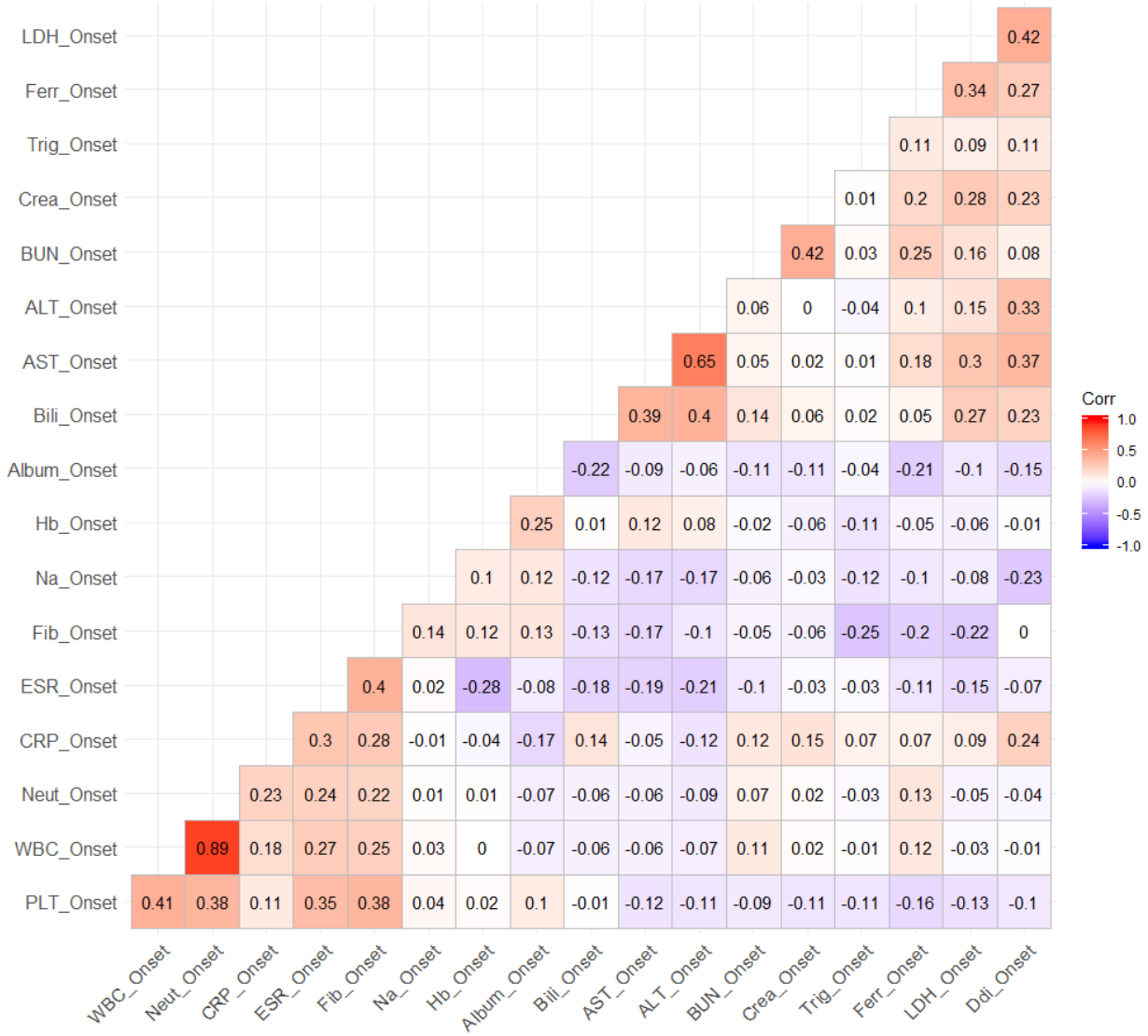


Fig. 1 – Correlation analysis for avaible laboratory parameters

We estimated one- through ten-class solutions using maximum likelihood estimation with robust standard errors. As LCA is an unsupervised clustering tool, classification was conducted without any consideration of clinical outcomes. Models with one to ten classes were estimated and compared. The best-fitting model was selected based on likelihood-based statistics (Akaike information criterion or AIC, Bayesian information criterion or BIC, R2-entropy) and model interpretability (i.e., face validity; classes including at least 5 % of cases). To assess the relation

between the obtained clusters and demographic features, clinical manifestation at the onset and outcome, a LCA regression analysis was performed using a three-step approach, which involves regressing class membership probabilities on the predictor/outcome variables while adjusting for the uncertainty in class assignments (Asparouhov & Muthén, 2014; Bakk et al., 2013; Vermunt, 2010)[23]. The analysis was performed using Latent Gold software 5.0 (Statistical Innovations, Belmont, MA). Preprocessing was done using R.

3. Results

3.1 Latent-Class Modeling: Identification of Number of Subgroups

A six-class model emerged as the best fitting solution according to selection criteria described above**.** BIC values improved as model complexity increased from 1 through 6 classes. Table 1 displays a summary of the likelihood-based statistics for each model. The selected model showed an R^2-entropy of 0.91, indicating an excellent discrimination between the estimated classes.

Tin contrast to BIC, AIC and AIC3 were lower in the seven-class model than in the six-class one; these last indexes are known to favour more complex models and overestimate the number of classes [24].

[23] Vermunt JK (2010). Latent class modeling with covariates: Two improved three-step approaches. Political analysis, 18, 450–469.
[24] Nylund, Karen L., Tihomir Asparouhov, and Bengt O. Muthén. "Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study." Structural equation modeling: A multidisciplinary Journal 14.4 (2007): 535-569.

To directly compare the two models, we performed a Parametric Bootstrapped Likelihood Ratio Test (BLRT) that suggested a significant improvement in fit for the seven-class model. Despite these elements supporting a better approximation of the data by this model, considering the lower entropy value for the seven-class solution, the small number of subjects in the seventh class (n = 18) and the similarity between the fifth and seventh class, the six-class model was selected. Tab. 1 summarizes the statistic fix indexes for each estimated model.

	LL	BIC(LL)	Npar	Class.Err.
1-Cluster	-31171,7850	62532,1026	32	0,0000
2-Cluster	-29828,0610	60039,0789	65	0,0134
3-Cluster	-29426,7484	59430,8779	98	0,0367
4-Cluster	-29175,1395	59122,0845	131	0,0453
5-Cluster	-29012,4576	58991,1448	164	0,0502
6-Cluster	-28891,2546	58943,1631	197	0,0540
7-Cluster	-28814,7112	58984,5005	230	0,0673
8-Cluster	-28744,7842	59039,0708	263	0,0634
9-Cluster	-28630,1751	59004,2768	296	0,0592
10-Cluster	-28592,0015	59122,3539	329	0,0529

Tab 1 – Statistical indexes for each estimated LCA model, from 1 to 10 classes

The identified classes differed significantly regarding the extent and combination of whole blood cells, platelets, ERS, CRP, AST, bilirubin, fibrinogen, d-dimer, ferritin, BUN and creatinine alterations. Hb resulted the least discriminating parameter between classes, showing a p of 0.23 and an R2 of 0,035; nevertheless, considered the prognostic value of this biomarker suggested by previous analysis[25] and the improvement in discrimination obtained through its inclusion, the Hb parameter was retained in the model (Table 2)

[25] Minoia, Francesca, et al. "Dissecting the heterogeneity of macrophage activation syndrome complicating systemic juvenile idiopathic arthritis." The Journal of rheumatology 42.6 (2015): 994-1001.

Parameter	*Cluster1*	*Cluster2*	*Cluster3*	*Cluster4*	*Cluster5*	*Cluster6*	*Wald*	*p-value*	*R^2*
Hb	-0,5378	-0,1477	0,4389	0,0244	-0,4568	0,6791	6,8724	0,23	0,0359
WBC	-1,1545	6,1904	-7,2768	-4,8663	9,3386	-2,2314	82,8956	2,10E-16	0,2457
PLT	-81,2539	197,3374	-58,826	-34,9133	36,7261	-59,0703	65,2151	1,00E-12	0,3021
ESR	4,3364	35,4148	-15,7138	-3,226	9,7773	-30,5886	136,3689	1,10E-27	0,2771
CRP	-1,1659	2,5068	-8,3952	-1,4368	5,3105	3,1806	68,4272	2,20E-13	0,144
AST	-503,298	-616,235	-559,028	-14,818	-434,31	2127,689	58,9198	2,00E-11	0,4059
LDH	-876,969	-2451,38	-2139,93	-1273,82	2365,071	4377,025	103,6368	9,00E-21	0,2997
Triglycerides	44,6517	-82,572	-36,2845	-31,2097	79,2386	26,1759	40,2388	1,30E-07	0,11
Na	0,0035	1,7811	1,3402	-0,5254	0,1619	-2,7613	24,8505	0,00015	0,0668
Bilirubin	-1,3767	-1,48	-1,2657	1,4188	-0,5242	3,2278	90,5862	5,10E-18	0,4113
Albumine	-0,086	0,1606	0,4158	-0,11	0,0302	-0,4105	35,0337	1,50E-06	0,1134
Fibrinogen	-37,8013	217,7495	-50,9048	-12,06	-35,3356	-81,6478	77,6321	2,60E-15	0,2948
Ddimer	-1079,05	-3333,95	-5876,52	-2424,41	1242,02	11471,91	88,8293	1,20E-17	0,284
Ferritin	3475,197	-11571,7	-11854,1	-9879,81	13570,74	16259,66	102,8296	1,30E-20	0,3257
BUN	-11,118	-13,3073	-12,8979	-5,0995	35,8342	6,5884	25,9253	9,20E-05	0,2176
Creatinine	-0,2526	-0,401	-0,2903	-0,2946	1,0939	0,1447	46,2566	8,00E-09	0,3624

Tab. 2 – Wald tests, p-values and R2 for the defining parameters (measured at MAS onset) of the latent clusters.

In the following paragraph we provide a brief description and a proposed sub-phenotype name based on the mean laboratory values of each class. Table 3 reports the mean values of laboratory parameters for each cluster.

Cluster	Hb	WBC	PLT	ESR	CRP	AST	LDH	Triglicerid	Na	Bilirubin	Albumin	Fibrinogen	Didimer	Ferritin	BUN	Creatinine
1	9,318	11,8643	129,2321	51,2879	12,1786	207,9481	2407,055	314,4136	134,7611	0,6282	3,0119	255,4234	6887,332	19327,8	26,3281	0,61
2	9,7081	19,2093	407,8234	82,3663	15,8512	95,0114	832,6474	187,19	136,5387	0,5249	3,2585	510,9743	4632,435	4280,933	24,1388	0,4615
3	10,2947	5,7421	151,66	31,2376	4,9492	152,2178	1144,095	233,4775	136,0978	0,7392	3,5137	242,32	2089,865	3998,489	24,5482	0,5722
4	9,8802	8,1526	175,5727	43,7254	11,9077	696,428	2010,202	238,5522	134,2323	3,4237	2,9878	281,1648	5541,976	5972,79	32,3466	0,568
5	9,399	22,3574	247,2122	56,7288	18,655	276,9358	5649,095	349,0005	134,9195	1,4807	3,1281	257,8891	9208,405	29423,35	73,2803	1,9564
6	10,5349	10,7874	151,4158	16,3629	16,5251	2838,935	7661,048	295,9378	131,9963	5,2327	2,6873	211,577	19438,3	32112,27	44,0345	1,0072

Tab 3 – Means value for parameters included in the LCA model characterizing each cluster. All values are converted to the international standard unit system: Hemoglobin, g/dl, White blood cell count, × 109/l Platelet count, × 109/l, ESR, mm/h, CRP, mg/dl, AST, U/l, LDH, U/l, Triglycerides,

mg/dl, Albumin, g/dl, Bilirubin, g/dl, Fibrinogen, mg/dl, Ferritin, ng/ml, D-dimer, ng/ml, BUN, mg/dl, Creatinine, mg/dl.

3.2 Laboratory and Clinical features of Subgroups

Cluster 1: Hyperferritinemia + thrombocytopenia

This is the largest subgroup, including 93 patients. These subjects presented with high ferritin at onset, with mean values around 20.000 ng/ml, associated with a low PLT count (mean value around 130.000/mm3). Inflammatory indexes showed high mean values in this group (ESR 51 mm/h, CRP 12,1 mg/dl). While transaminases were around four times above the norm, no other signs of organ dysfunction characterized this cluster.

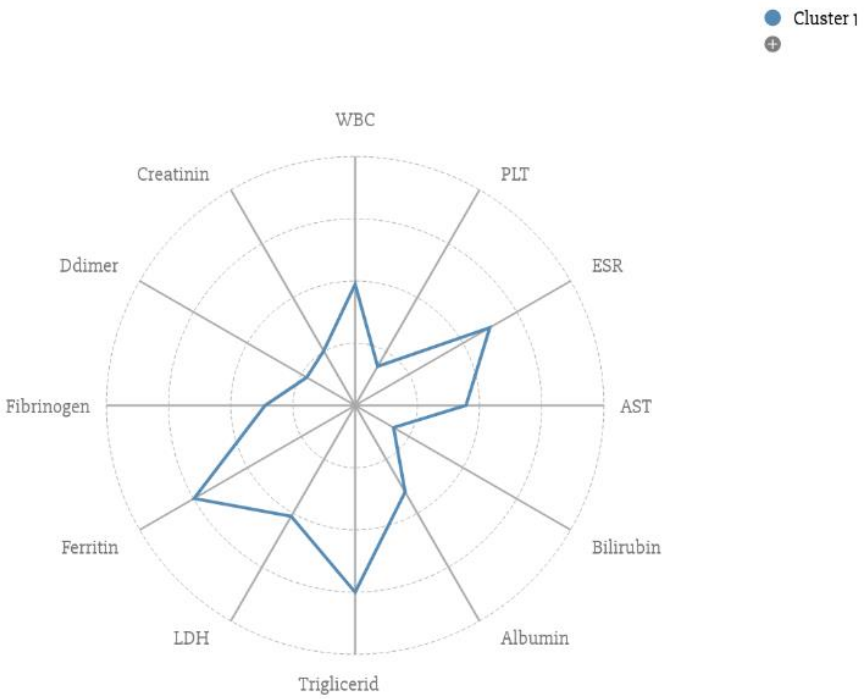


Fig 2- Cluster 1

Cluster 2: Hyper-inflammation

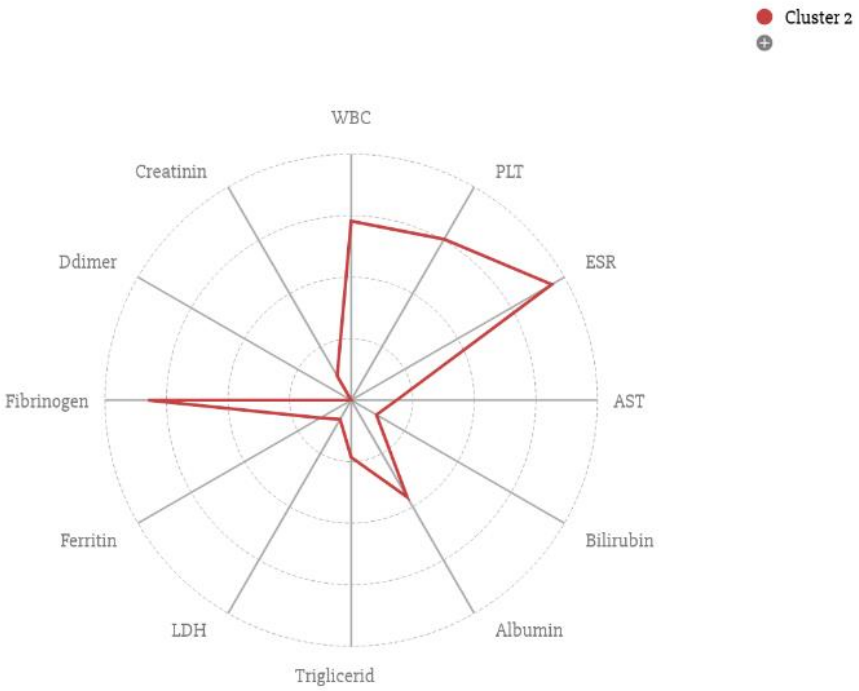


Fig 3- Cluster 2

This cluster containing 83 patients, was almost exclusively characterized by very high inflammatory indexes (ESR, CRP, Fibrinogen), PLT and WBC values, with relatively low ferritin values (mean values of around 4000 ng/dl) and a very low probability of alteration of parameters measuring organ damage. Among this group, the lowest values of serum aspartate aminotransferase (AST) and Creatine were observed (respectively mean values of 95 U/l and 0,4 mg/dl). These patients tended to be younger and with lower SJIA duration at MAS onset then other subjects.

Cluster 3: Pancytopenia

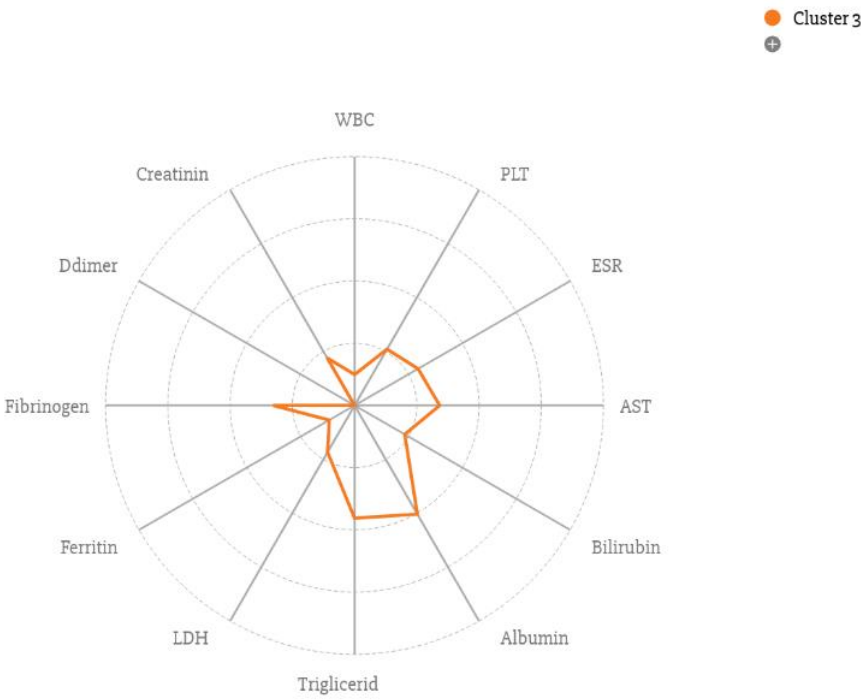


Fig 4- Cluster 3

This subgroup (n. = 82) presented with hematologic involvement, with the lowest values of WBC and PLT, but with few other organ dysfunctions. Moreover, these patients exhibited the lowest values of CRP (mean around 5 mg/dl) and ferritin levels similar to cluster 2. Modest hepatic involvement (AST around 152 U/l, Bilirubin 0,7 g/dl) was also present in this group.

Cluster 4: Hepatobiliary dysfunction

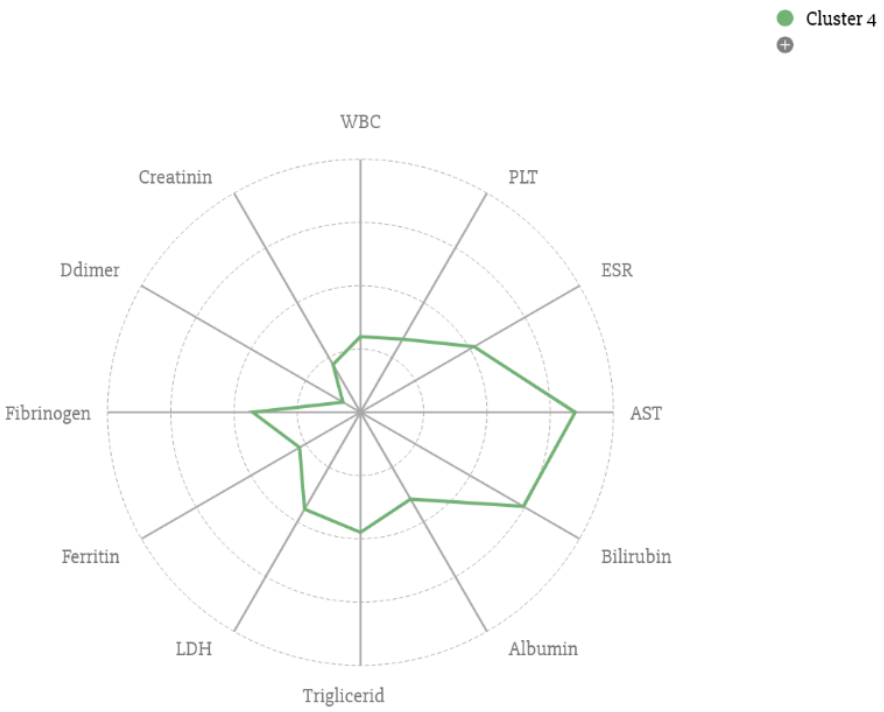


Fig 5- Cluster 4

This group (n. = 56) was mainly characterized by significant hepatic and biliary involvement, in absence of coagulopathy and renal dysfunction. Mean values of AST were around 700 U/L. These patients exhibited the highest likelihood of Jaundice, with a mean Bilirubin value of 3,4 mg/dl. Albumin levels were slightly reduced; however, no signs of liver insufficiency were evident in this group.

Cluster 5: Acute Kidney Injury + High LDH, Ferritin, Triglycerides

Fig 5- Cluster 4

Fig 5- Cluster 4

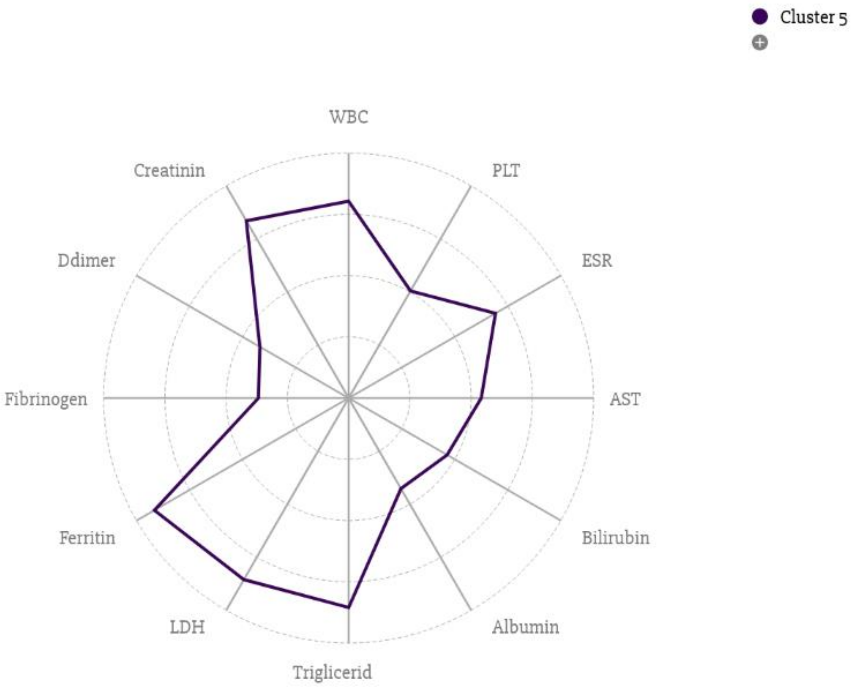


Fig 6- Cluster 5

These patients (n.=26) exhibited the highest levels of creatinine and BUN, suggesting acute kidney injury. The profile was also characterized by Ferritin and LDH levels among the highest, coupled with the highest levels of triglycerides. WBC and CRP also reached the highest levels in this group, associated with a relatively high platelet count (around 250.000/mm3). Compared to cluster 4, the hepatobiliary dysfunction was milder, with AST mean levels around 250 U/l.

Cluster 6: Liver Insufficiency + Coagulopathy

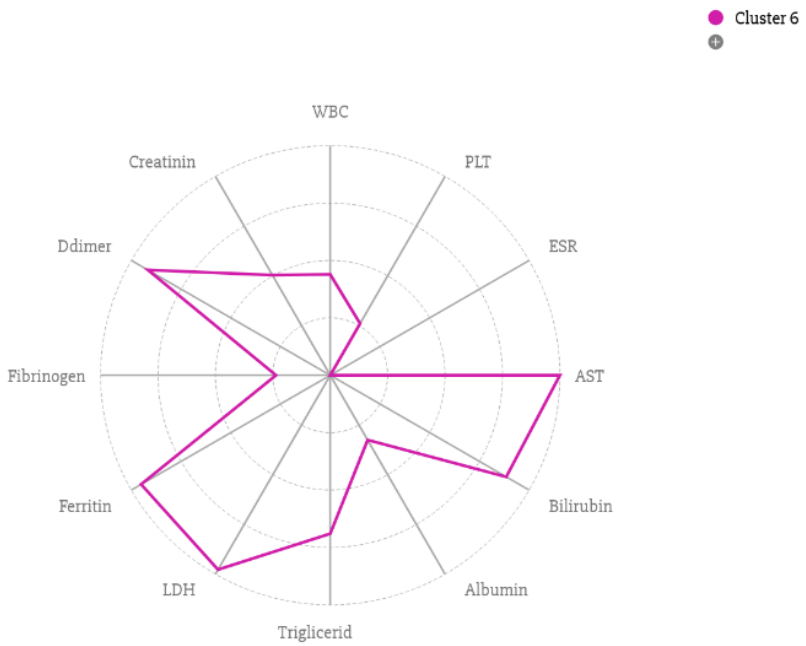


Fig 7- Cluster 6

This cluster (n.=22) included patients with evidence of hepatic failure, as revealed by mean values of AST around 2800 U/L, mean Bilirubin levels around 5 mg/dl and mean Albumin values of 2,6 g/dl. The lowest value of ESR and Fibrinogen also indicate severe coagulopathy. The highest levels of Ferritin, LDH, D-Dimer were observed in this group. There was also evidence of kidney involvement (mean Creatinine value around 1 mg/dl), albeit to a lower extent than cluster 5.

A comparison of the laboratory profiles of the different clusters is provided in Fig. 8

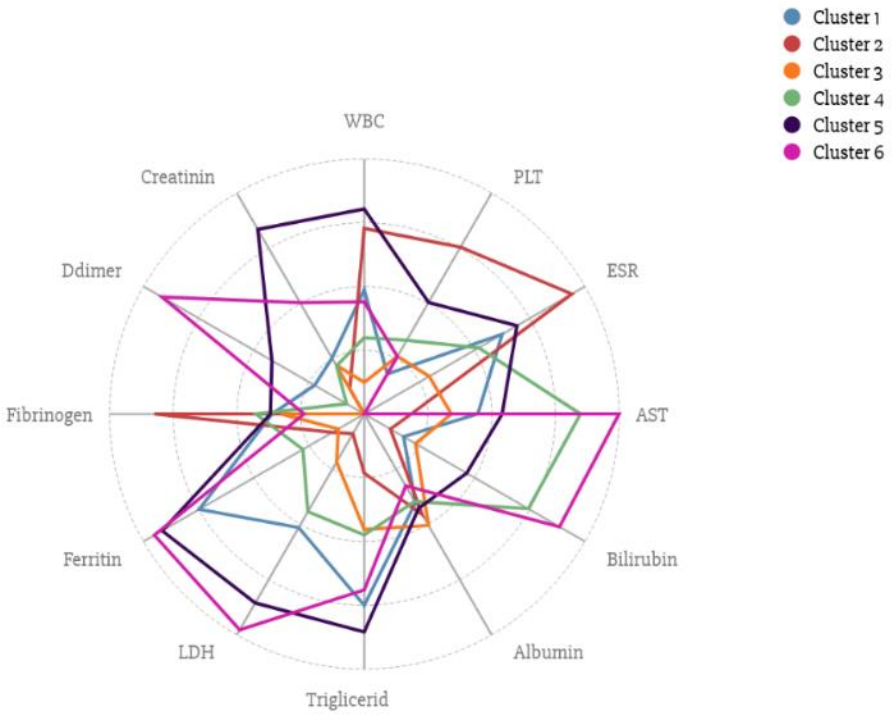


Fig 8- All clusters

3.3 Comparison of outcomes and predictor variables between clusters

We found substantial differences between the clusters regarding demographic features, treatment requirement and clinical outcomes.

Patients in Cluster 2 (OR 0.93, 95%Cl 0.93 - 0.94, p = 0.0025) and Cluster 1 (0.96, 95%Cl 0.95 - 0.96) were younger at MAS onset, while Cluster 3 members were the eldest (OR 1.06, 95%Cl 1.05 – 1.07). (Cluster 6 OR 1.01, 95%Cl 1.00 – 1.01)

SJIA duration at onset was lower for cluster 6 (OR 0.90, 95%Cl 0.89 – 0.91, 0 = 0.0034) and Cluster 2 (OR 0.92, 95%Cl 0.91 – 0.93), while Cluster 3 members had the longest disease duration (OR 1.14, 95%Cl 1.12 - 1.15). No association was found between age at SJIA onset and cluster membership.

No significant associations were observed between the clusters and geographic provenience, specialty of the referral Unit and gender.

Subjects in Cluster 5 exhibited the worst clinical outcomes, including a significant higher mortality rate (OR 2.03, 95%Cl 2.27 - 1.82, p = 0.029).

Infact, Cluster 5 showed the highest likehood of severe course (OR 3.61, 95%Cl 3.97 - 3.28; p = 0.0001). Cluster 6 presented with a OR 1.25, (95%Cl 1.27 - 1.23) Subjects in Cluster 2 had the lowest probability of severe course (OR 0.34, 95%Cl 0.449 - 0.055; p = 0.0001).

Cluster 2 membership was also protective for fatal outcome (OR 0.52, 95%Cl 0.58 - 0.47, p = 0.029).

Cluster 5 (OR 2.76, 95%Cl 2.95 - 2.59, p<0.0001) and Cluster 6 (OR 1.05, 95%Cl 1.05 – 1.05) had the highest risk of multiple organ failure. Cluster 2 membership resulted as protective of multiple organ failure (OR 0.66, 95%Cl 0.68 – 0.64).

Furthermore, Cluster 5 patients had the highest probability of developing CSN dysfunction (OR 2.54, 95%Cl 2.68 - 2.41, p <0.001), followed by Cluster 6 (OR 1.41, 95%Cl 1.44 - 1.38). Cluster 2 had the lowest probability of CSN involvement (OR 0.55, 95%Cl 0.53 - 0.56).

Likewise, the highest rate of hemorrhagic complications occurred in Cluster 5 patients (OR 1.72, 95%Cl 1.82 - 1.62, p = 0.003), followed by cluster 6 (OR 1.33, 95%Cl 1.37 - 1.29). The lowest probability was observed in Cluster 2 (OR 0.63, 95%Cl 0.57 - 0.67) and Cluster 3 (OR 0.79, 95%Cl 0.77 - 0.80).

Unsurprisingly, the use of corticosteroids and others immunosuppressors was more frequent for clusters 4, 5 and 6 than for clusters 1, 2 and 3. Cluster 5 (OR 1.66, 95%Cl 1.79 – 1.54, p 0.021) and Cluster 4 (OR 1.52, 95%Cl 1.62 – 1.43) had the highest probability to be treated with Etoposide. Cluster 3 had the lower probability of Etoposide treatment (OR 0.49, 95%Cl 0.45 – 0.55). No significant differences were detected regarding the use of cyclosporine, intravenous immune globulins, plasmapheresis and biologic agents. No significant interaction between the six classes and treatment assignment in predicting severe course and mortality was detected.

4. Discussion

In this work, LCA identified distinct subgroup with different baseline biomarkers profiles, demographic features and clinical outcomes in a large, multinational cohort of sJIA-MAS patients. The high number of revealed clusters, together with the heterogeneity of their associations, confirm the complexity of this multifaceted condition. Despite the elaborateness of the model, the clusters represent recognizable patterns of organ involvement that are known to characterize HLH/MAS, showing a good face validity from a clinical point of view. These subgroups seem unrelated to the geographic provenance of patients and the specialty of the referral center, apparently reflecting biological and clinical differences. Among the defining laboratory parameters, Hb levels resulted the less discriminative variable, despite being identified as

prognostic markers by other analysis. We found, instead, that biomarkers expressing hepatobiliary dysfunction, namely transaminases and bilirubin, contributed more to cluster definition, followed by creatinine, ferritin, and platelet count. In other words, these parameters showed the greatest variability among MAS patients. Exploring the prognostic utility of these biomarkers in further studies could improve the prediction of outcomes. The emerged picture reveals a wide spectrum of severity encompassing/ranging from few laboratory alterations and lower probability of organ damage, as in cluster 1, 2 and 3, to severe complications and multiple dysfunctions, which occur in smaller subgroup, such as cluster 5 and 6. Most of MAS patients seem to show hyperferritinemia in combination with thrombocytopenia (cluster 1) or elevated inflammatory indexes (cluster 2). Cluster 2, in particular, showed a markedly lower risk of CSN, hemorrhagic, hepatic, renal complications, ICU admission and mortality. Infact, this profile differed from a typical, uncomplicated SJIA presentation only regarding the level of ferritin (mean values around 4600 ng/ml). These patients were younger and were among those with lower sJIA duration at MAS onset. This finding agrees is in line with a previous multivariate analysis of the same dataset, which identified an older age at MAS onset as a negative prognostic factor. We argue that this result in the original analysis was driven by the presence of this young, low-risk subgroup, rather than a causal impact of an older age on the clinical outcomes. Moreover, disease duration was also relatively lower for patients in the high-risk Cluster 6. Therefore, the association between low sJIA duration and could simply reflect difficulties differential diagnosis between an active, uncomplicated sJIA at the onset and MAS. Similar to cluster 2, the group defined by isolated hematologic alterations, namely cluster 3, had low probability of developing organ involvement. These subjects were older and developed MAS after a longer disease

duration compared to other clusters. Among the clusters with high incidence of complications, cluster 4, which was characterized by hepatobiliary dysfunction but lower probability of coagulopathy and renal dysfunction, showed the mildest course. A minority of patients were classified in the 2 groups with the most severe presentation. The worst outcome was associated with Cluster 5; these patients already presented at onset with acute kidney injury, coagulopathy with high D-Dimer levels despite milder hepatic involvement, high LDH and inflammatory indexes. This profile was associated with the highest rate of CSN involvement, hemorrhagic complications, multiple organ failure, severe course and death. Understandably, these patients received more aggressive treatments, including etoposide. The combination of marked LDH and D-dimer alteration with high creatinine levels in this group could suggest the occurrence of a well-recognized complication of HLH/MAS, namely thrombotic thrombocytopenic purpura (TTP). The elevation of Bilirubin despite the relatively mild hepatitis could also reflect a hemolytic process, although Hb levels were not significantly lower in this cluster (mean values of 10,5 g/dl). However, unlike in classical thrombotic microangiopathy presentations, this profile was not associated with significant thrombocytopenia. Rather, this pattern resembles a form of severe DIC-like coagulative dysfunction leading to acute kidney injury (AKA). The other high-risk subgroup, Cluster 6, showed a full-blown MAS presentation with severe pancytopenia, hepatobiliary dysfunction up to liver failure, coagulopathy and renal involvement, albeit in absence of AKA. This group showed a slightly lower probability of adverse outcomes (CNS, Hemorrhagic, MOF, severe course, death) than to cluster 5; however, direct paired comparison between the two clusters failed to demonstrate a significant difference, except for the occurrence multiple organ failure ($p = 0.02$). Interestingly, these patients had also lower disease

duration at MAS onset, suggesting that some sJIA patients are actually at risk of severe MAS early in their clinical course. Apart from disease duration, we couldn't find any baseline clinical or demographic factor able to discriminate the two high-risk subgroups from other MAS patients. Overall, these clusters recapitulate the full spectrum of known MAS manifestations and could represent different stages of the pathophysiological process leading from hyperinflammation to single organ damage to multiple organ failure. The current understanding of sJIA-MAS pathogenesis, supported by accumulating evidence in animal and human studies, suggests a causal role for a cytokine feedback loop involving IL-18 and IFNγ. Serum IL-18 concentrations are elevated in a subgroup of patients with SJIA and adult-onset Still's disease[26], and their increment precede and correlate with ferritin, transaminase and CRP levels[27]. Moreover, high IL-18 levels predispose patients with SJIA - and several other autoinflammatory conditions[28] - to develop MAS [29]. The mechanism underlying this process seems to involve the specific activation of the IFNγ pathway in T cells and macrophages by excessive Il-18 signaling[30]. The pivotal role of IFNγ in HLH/MAS has been demonstrated in several neutralization and clinical studies, which showed a strict correlation between levels of IFNγ -induced chemokines in tissues and organ damage[31]. In

[26] Colafrancesco, Serena, et al. "IL-18 serum level in adult onset Still's disease: a marker of disease activity." International journal of inflammation 2012 (2012).
[27] Shimizu, Masaki, et al. "Distinct cytokine profiles of systemic-onset juvenile idiopathic arthritis-associated macrophage activation syndrome with particular emphasis on the role of interleukin-18 in its pathogenesis." Rheumatology 49.9 (2010): 1645-1653.
[28] Canna SW, de Jesus AA, Gouni S, et al. An activating NLRC4 inflammasome mutation causes autoinflammation with recurrent macrophage activation syndrome. Nat Genet. 2014;46:1140–6
[29] Shimizu M, Nakagishi Y, Inoue N, Mizuta M, Ko G, Saikawa Y, et al. Interleukin-18 for predicting the development of macrophage activation syndrome in systemic juvenile idiopathic arthritis. Clin Immunol. 2015;160(2):277–81
[30] Girard-Guyonvarc'h, C., Palomo, J., Martin, P., Rodriguez, E., Troccaz, S., Palmer, G., & Gabay, C. (2018). Unopposed IL-18 signaling leads to severe TLR9-induced macrophage activation syndrome in mice. Blood, 131(13), 1430-1441.
[31] Bracaglia, C., de Graaf, K., Marafon, D. P., Guilhot, F., Ferlin, W., Prencipe, G., ... & Grom, A. A. (2017). Elevated circulating levels of interferon-γ and interferon-γ-induced chemokines characterise patients with macrophage

this view, the dysregulation of the IL-18-IFNγ feedback underlies a continuum of clinical and laboratory manifestations that encompasses both sJIA and MAS as two extremes of a spectrum, rather than as two distinct diseases[32]. Different phases of this immunopathogenic dynamics can produce distinct phenotypes, including the occurrence of typical MAS biomarker and pathological abnormalities in absence of clinical complications, sometimes referred at as "occult MAS"[33]. Murine models of secondary HLH[34] [35] support the notion that the IL-18 induced production of IFNγ begins at the organ level, i.e. in liver and bone marrow tissues, where it drives inflammation and tissue damage; as already proposed for other systemic hyperinflammatory conditions[36], when the extent or duration of the inflammatory process reaches a certain threshold, the response will spill over into other compartments, disseminating systemically and ultimately leading to multiple organ failure. In the light of this pathogenetic model, we argue that the identified clusters present with different severity and patterns of organ involvement because of different stages of this pathogenetic process. Cluster 2 could represent a subset of patients that resides at the sJIA edge of the sJIA-MAS spectrum, and therefore have a low probability to develop a full MAS clinical pictures. According to avaible evidences, this profile could be driven

activation syndrome complicating systemic juvenile idiopathic arthritis. Annals of the rheumatic diseases, 76(1), 166-172.

[32] Bracaglia, C., Prencipe, G., & De Benedetti, F. (2017). Macrophage activation syndrome: different mechanisms leading to a one clinical syndrome. Pediatric Rheumatology, 15(1), 5.

[33] Behrens EM, Beukelman T, Paessler M, et al. Occult macrophage activation syndrome in patients with systemic Juvenile idiopathic arthritis. J Rheumatol. 2007 May;34(5):1133–1138. PubMed PMID: 17343315

[34] Reinhardt, R. L., Liang, H. E., Bao, K., Price, A. E., Mohrs, M., Ben, L. K., & Locksley, R. M. (2015). A Novel Model for IFN-g–Mediated Autoinflammatory Syndromes. 1401992.

[35] Buatois, V., Chatel, L., Cons, L., Lory, S., Richard, F., Guilhot, F., ... & Kosco-Vilbois, M. H. (2017). Use of a mouse model to identify a blood biomarker for IFNγ activity in pediatric secondary hemophagocytic lymphohistiocytosis. Translational Research, 180, 37-52.

[36] Dick, T. E., Molkov, Y., Nieman, G., Hsieh, Y. H., Jacono, F. J., Doyle, J., ... & Vodovotz, Y. (2012). Linking inflammation, cardiorespiratory variability, and neural control in acute inflammation via computational modeling. Frontiers in physiology, 3, 222.

by elevated serum levels of IL-18 which have not yet determined a widespread induction of the IFNγ pathway. Cluster 1, showing some early laboratory alteration such as hyperferritinemia and thrombocytopenia, could capture the transition from IL-18 to IFNγ hyperproduction, in other words from SJIA to MAS. Cluster 3 and Cluster 4, which are characterized by isolate organ involvement, could suggest an hyperproduction of IFNγ that is confined to the organ levels. Finally, Cluster 6 and 5 include patients that already present with systemic hyperinflammation and severe organ damage at onset. Although evidences suggest that the transition from one of these phases to the others is often non-linear and threshold-wise, these findings highlight the importance of a prompt diagnosis and early treatment start for MAS prognosis. Our results must be interpreted with caution, considering the many limitations of this study. We analyzed a retrospective dataset collected through a web-based survey. Significant percentage of missing values were present for many variables, requiring imputation. The original study was designed to define a new set of diagnostic criteria for sJIA-MAS, with no original intention to investigate prognostic factors. As a consequence, many important prognostic information, such as the time of occurrence of complications and outcomes and duration of treatments, were non detailed and unavailable for the analysis. Multiple biomarkers assessments at pre-defined time periods were not collected, foreclosing a proper longitudinal analysis. Due to the lack of temporal details and the observational nature of the dataset, we couldn't account for treatment use and other post-onset variables when investigating the relationships between clusters and outcomes. Moreover, in absence of diagnostic gold standards, study inclusion was based on the diagnostic judging of the referral physicians, inherently introducing some degree of subjectivity and circularity. Furthermore, the generability of our finding is limited by the fact that the study

population included MAS cases diagnosed from 2002 to 2012; consequently, relatively few patients were treated with the now more commonly employed biologic therapies such as Anti-IL6 and Anti-IL1 agents, which recently proved their efficacy in both sJIA and MAS and could possibly modify the clinical and laboratory MAS presentation. Finally, due to the small simple size of some clusters, we weren't able to compare the additional information provided by the latent classes with respect to the prognostic value of single laboratory parameters, and assess if the improvements is substantial enough to support the use of this classification as a prognostic tool. Further research on external datasets are warranted to investigate the validity and prognostic and predictive utility of these subgroups. Despite these limitations and the inherently explorative nature of our analysis, we believe that these findings may add to the understanding of the complexity of MAS and provide a proof-of-concept for MAS patient stratification based on routine avaible biomarkers. Among the strengths of the study are the analysis of the largest multinational MAS cohort collected so far, and the use of a patient-centered analytic methodology which, providing a data-drive classification that did not include clinical outcomes as class-defining variables, allow to estimate the unbiased associations between subgroups and outcomes.

5. Conclusions

In this explorative analysis we identified different subgroups of sJIA-MAS patients with distinctives laboratory profiles at disease onset. The clusters show different clinical and demographic features clinical outcomes, and treatment requirement. These findings provide insights on the heterogeneity of MAS severity and manifestations and highlight the prognostic

values of laboratory patterns at onset. Further analyses are needed to validate these results and investigate the potential predictive value of the identified clusters.

References

- Behrens EM, Beukelman T, Paessler M, et al. Occult macrophage activation syndrome in patients with systemic Juvenile idiopathic arthritis. J Rheumatol. 2007 May;34(5):1133–1138. PubMed PMID: 17343315
- Berlin, Kristoffer S., Natalie A. Williams, and Gilbert R. Parra. "An introduction to latent variable mixture modeling (part 1): Overview and cross-sectional latent class and latent profile analyses." Journal of Pediatric Psychology 39.2 (2014): 174-187.
- Bracaglia, C., de Graaf, K., Marafon, D. P., Guilhot, F., Ferlin, W., Prencipe, G., ... & Grom, A. A. (2017). Elevated circulating levels of interferon-γ and interferon-γ-induced chemokines characterise patients with macrophage activation syndrome complicating systemic juvenile idiopathic arthritis. Annals of the rheumatic diseases, 76(1), 166-172.
- Bracaglia, C., Prencipe, G., & De Benedetti, F. (2017). Macrophage activation syndrome: different mechanisms leading to a one clinical syndrome. Pediatric Rheumatology, 15(1), 5.
- Bracaglia, Claudia, Giusi Prencipe, and Fabrizio De Benedetti. "Macrophage activation syndrome: different mechanisms leading to a one clinical syndrome." Pediatric Rheumatology 15.1 (2017): 5.
- Buatois, V., Chatel, L., Cons, L., Lory, S., Richard, F., Guilhot, F., ... & Kosco-Vilbois, M. H. (2017). Use of a mouse model to identify a blood biomarker for IFNγ activity in pediatric secondary hemophagocytic lymphohistiocytosis. Translational Research, 180, 37-52.
- Campbell, Christina A., et al. "Assessing Intervention Needs of Juvenile Probationers: An Application of Latent Profile Analysis to a Risk–Need–Responsivity Assessment Model." Criminal Justice and Behavior 46.1 (2019): 82-100.
- Canna SW, de Jesus AA, Gouni S, et al. An activating NLRC4 inflammasome mutation causes autoinflammation with recurrent macrophage activation syndrome. Nat Genet. 2014;46:1140–6
- Colafrancesco, Serena, et al. "IL-18 serum level in adult onset Still's disease: a marker of disease activity." International journal of inflammation 2012 (2012).
- Crayne, Courtney B., et al. "The Immunology of Macrophage Activation Syndrome." Frontiers in immunology 10 (2019).
- Dick, T. E., Molkov, Y., Nieman, G., Hsieh, Y. H., Jacono, F. J., Doyle, J., ... & Vodovotz, Y. (2012). Linking inflammation, cardiorespiratory variability, and neural control in acute inflammation via computational modeling. Frontiers in physiology, 3, 222.
- Girard-Guyonvarc'h, C., Palomo, J., Martin, P., Rodriguez, E., Troccaz, S., Palmer, G., & Gabay, C. (2018). Unopposed IL-18 signaling leads to severe TLR9-induced macrophage activation syndrome in mice. Blood, 131(13), 1430-1441.

- Minoia, Francesca, et al. "Clinical features, treatment, and outcome of macrophage activation syndrome complicating systemic juvenile idiopathic arthritis: a multinational, multicenter study of 362 patients." Arthritis & Rheumatology 66.11 (2014): 3160-3169.
- Minoia, Francesca, et al. "Criteria for Cytokine Storm Syndromes." Cytokine Storm Syndrome. Springer, Cham, 2019. 61-79.
- Minoia, Francesca, et al. "Dissecting the heterogeneity of macrophage activation syndrome complicating systemic juvenile idiopathic arthritis." The Journal of rheumatology 42.6 (2015): 994-1001.
- Muthén, B. (2004). Latent variable analysis. In D. Kaplan (Ed.), The Sage handbook of quantitative methodology for the social sciences (pp. 345–368). Thousand Oaks, CA: Sage Publications.
- Nylund, Karen L., Tihomir Asparouhov, and Bengt O. Muthén. "Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study." Structural equation modeling: A multidisciplinary Journal 14.4 (2007): 535-569.
- Ravelli, Angelo, et al. "Macrophage activation syndrome." Hematology/Oncology Clinics 29.5 (2015): 927-941.
- Reinhardt, R. L., Liang, H. E., Bao, K., Price, A. E., Mohrs, M., Ben, L. K., & Locksley, R. M. (2015). A Novel Model for IFN-g–Mediated Autoinflammatory Syndromes. 1401992.
- Shimizu M, Nakagishi Y, Inoue N, Mizuta M, Ko G, Saikawa Y, et al. Interleukin-18 for predicting the development of macrophage activation syndrome in systemic juvenile idiopathic arthritis. Clin Immunol. 2015;160(2):277–81
- Shimizu, Masaki, et al. "Distinct cytokine profiles of systemic-onset juvenile idiopathic arthritis-associated macrophage activation syndrome with particular emphasis on the role of interleukin-18 in its pathogenesis." Rheumatology 49.9 (2010): 1645-1653.
- Vermunt JK (2010). Latent class modeling with covariates: Two improved three-step approaches. Political analysis, 18, 450–469.

Chapter 3 – Prognostic values of laboratory markers in sJIA-MAS

1. Introduction

Macrophage activation syndrome (MAS) is a term describing a constellation of clinical and laboratory abnormalities that can occur as a complication of many systemic diseases, with a particular frequency in systemic juvenile idiopathic arthritis[37], in which occurs in 7–13 % of patients. The disease, now recognized as part of the of hemophagocytic syndromes (HLH), is characterized by the uncontrolled release of proinflammatory mediators resulting in a widespread systemic inflammation[38]. The resulting clinical and laboratory picture encompasses a wide range of manifestation such as hyperferritinemia, cytopenias, liver dysfunction, and coagulopathy resembling disseminated intra-vascular coagulation and neurologic dysfunction. In a subgroup of patients, the hyperinflammatory process evolves in organ damage, multiorgan failure[39] and finally leads to fatal outcome, with reported mortality rates of 22–30 %[40]. Despite the severity of this complication, prognostic factors of MAS have not been fully investigated. In the patients with genetic forms of HLH, hyperferritinemia and cerebrospinal fluid (CSF) pleocytosis were has been reported to be independent risk factors of adverse outcome[41]. In

[37] Girard-Guyonvarc'h, C., Palomo, J., Martin, P., Rodriguez, E., Troccaz, S., Palmer, G., & Gabay, C. (2018). Unopposed IL-18 signaling leads to severe TLR9-induced macrophage activation syndrome in mice. Blood, 131(13), 1430-1441.

[38] Buatois, V., Chatel, L., Cons, L., Lory, S., Richard, F., Guilhot, F., ... & Kosco-Vilbois, M. H. (2017). Use of a mouse model to identify a blood biomarker for IFNγ activity in pediatric secondary hemophagocytic lymphohistiocytosis. Translational Research, 180, 37-52.

[39] Ravelli, Angelo, et al. "Macrophage activation syndrome." Hematology/Oncology Clinics 29.5 (2015): 927-941.

[40] Sawhney S, Woo P, Murray KJ. Macrophage activation syndrome: a potentially fatal complication of rheumatic disorders. Arch Dis Child. 2001; 85:421–6

[41] H. Trottestam, E. Berglof, A. Horne, E. Onelov, K. Beutel, K. Lehmberg, E. Sieni,

another study, hyperbilirubinemia, hypoalbuminemia, prolonged APTT and elevated LDH predicted early mortality; the authors concluded that biomarkers of liver function damage are the most useful prognostic factors[42]. In sJIA-related MAS, however, few evidences on factors predicting the development of a complicated course are avaible. The identification of prognostic factors could improve MAS outcome allowing for patient stratification and the rational choice of appropriate treatments for high-risk subgroups. In this study, we aim to investigate the prognostic value of routinely gathered laboratory biomarkers, identifying the most predictive parameters for the prediction of unfavorable outcomes, focusing on severe course (as defined as the occurrence of of intensive care admission or death) , mortality and the development of CNS involvement.

2. Methods

2.1 Study population

The sJIA-MAS cohort analyzed in this study has been already described in previous chapters and in dedicated papers[43]. As our emphasis here is on prediction of outcomes, we included in the analysis only laboratory parameters collected at onset and other demographic features, like age, that are avaible at MAS diagnosis. The "severe course" outcome is a composite outcome defined as the need of admission to the ICU or a death. CNS disease was defined as the occurrence of seizures, lethargy, irritability, headache, mood changes, confusion, coma or other manifestation.

T. Silfverberg, M. Arico, G. Janka, J.I. Henter, Risk factors for early death in children with haemophagocytic lymphohistiocytosis, Acta Paediatr. 101 (3) (2012)

[42] Tang, Y., & Luo, Z. (2018). Prognostic factors of early death in children with hemophagocytic lymphohistiocytosis. Cytokine, 110, 481-482. doi:10.1016/j.cyto.2017.09.023

[43] Minoia, Francesca, et al. "Dissecting the heterogeneity of macrophage activation syndrome complicating systemic juvenile idiopathic arthritis." The Journal of rheumatology 42.6 (2015): 994-1001.

As a substantial percentage of observation for some variable was missing (Fig. 1), data were imputed by multiple imputation using the R package MICE.

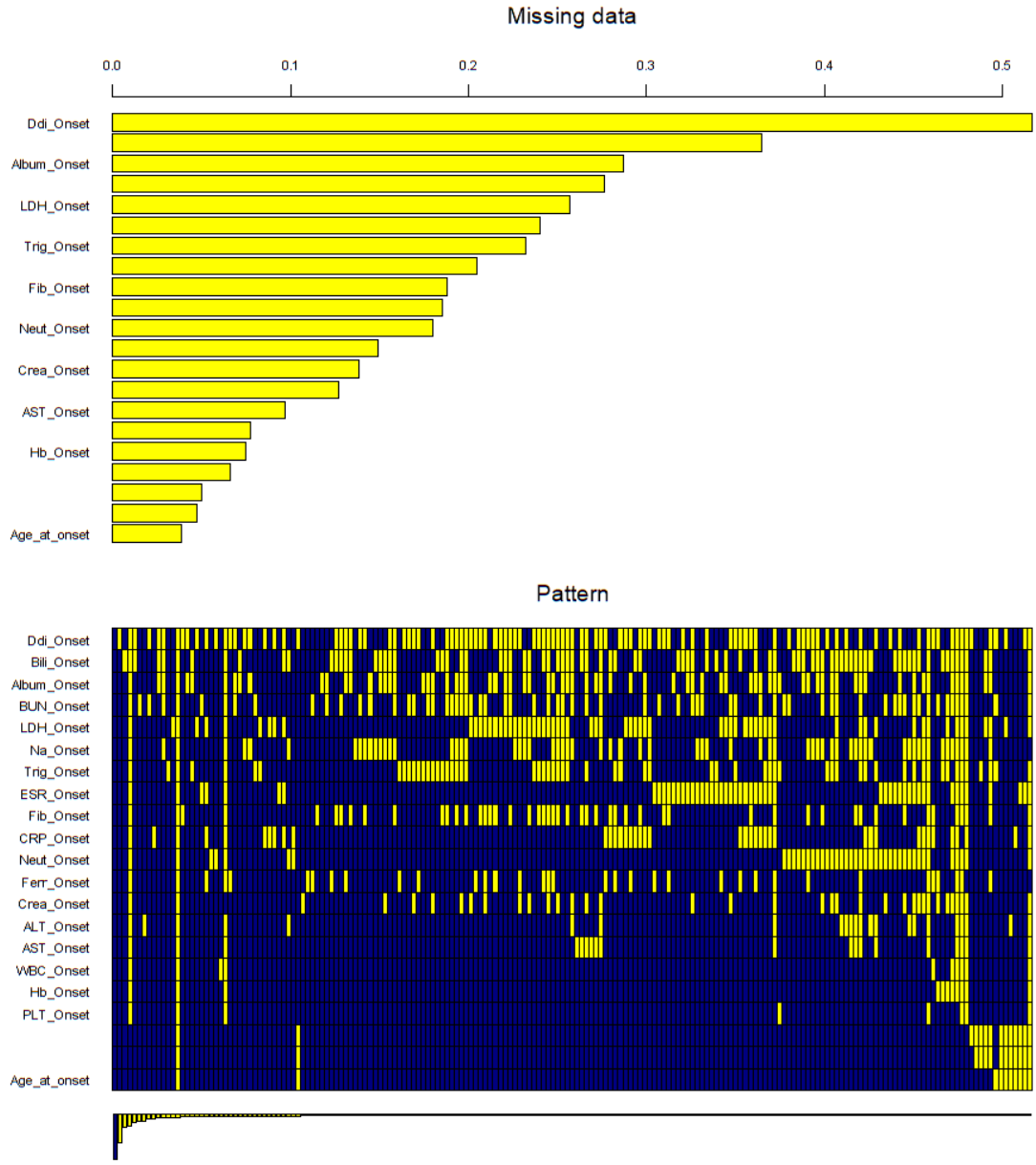


Fig. 1 - Percentage of missing data by variable

2.2 Data analysis

We applied multivariate regression analysis to investigate the relationships between predictors and outcomes. To account for the low rate of events for the outcomes of interest (35% of the sample developed CNS disease, 27% had a severe course and 8% died), we employed Firth's penalized likelihood in other to reducing small-sample bias in regression analysis[44]. To obtain unbiased estimation of the association between risk factors and outcomes, we used Direct Acyclic Graphs (DAGs) to select the set of variables to be included in the regression analysis. DAGs are the tools that allow to conceptualize the supposed causal relationships between variables according to expert knowledge or previous evidences. Visualizing the causal etiological network of involved factors, DAGs can be most useful to identify confounders to adjust for in the analysis, in this way minimizing bias[45]. Once the DAG is designed A mathematical ruleset, known as d-separation criteria, allows to select the variables to be included in the regression model without the risk of over-adjusting[46]. We thus constructed a DAG of hypothesized causal relationships between the factors driving outcomes in sJIA-MAS to identify the initial set of variables to be included in the model. The regression model was then implemented applying the Allen–Cady Modified Backward Selection procedure. This methodology consists in ranking by importance the candidate variables that are not forced into the model due to their a priori interest or known confounding role; then, ranked variables are deleted through a backward selection, in order of ascending importance. The procedure

[44] Heinze, Georg, et al. "Firth's bias reduced logistic regression." *R package version* 1 (2013).
[45] Williams, Thomas C., et al. "Directed acyclic graphs: a tool for causal studies in paediatrics." *Pediatric research* 84.4 (2018): 487.
[46] Snowden, Jonathan M., et al. "The Curse of the Perinatal Epidemiologist: Inferring Causation Amidst Selection." *Current epidemiology reports* 5.4 (2018): 379-387.

is advocated for by experts to limit false-positive results, due to univocally pre-specified sequence of model [47]. Finally, to detect possible informative cut-offs values useful to predict the outcomes, we employed Classification and Regression Trees (CART) analysis

CARTs are predictive algorithms that perform a binary, recursive splitting of observations to obtain subgroups with a similar response[48]. CARTs can be therefore used identify cut-off thresholds for the predictors and rank relevant risk factors. Thus, they are a useful for constructing decision trees to support decision-making in clinical practice[49]. CARTs were implemented through the R package party. ROC curve then was used to assess the accuracy after 10-fold cross-validation. Penalized likehood regression models were implemented in STATA 13 (Stata Corporation, College Station, Texas).

3. Results

Based on the DAGs, we included in the model demographic features such as age at MAS onset, age at sJIA onset and sJIA duration at MAS, as well as laboratory parameters at onset. In highly correlated couple of variables (WBC and neutrophilic count, AST and ALT, creatinine and BUN) only one of the two parameters was included, based on previous knowledge. In the following sections we describe the results for each outcome.

3.1 Severe course

[47] Vittinghoff, Eric, et al. *Regression methods in biostatistics: linear, logistic, survival, and repeated measures models*. Springer Science & Business Media, 2011.
[48] Loh, Wei-Yin. "Classification and regression tree methods." *Wiley StatsRef: Statistics Reference Online* (2014).
[49] Mburu, Josephine W., et al. "Use of classification and regression tree (CART), to identify hemoglobin A1C (HbA1C) cut-off thresholds predictive of poor tuberculosis treatment outcomes and associated risk factors." *Journal of Clinical Tuberculosis and Other Mycobacterial Diseases* 11 (2018): 10-16.

Results from the final model for the prediction of severe course are reported in Tab. 1

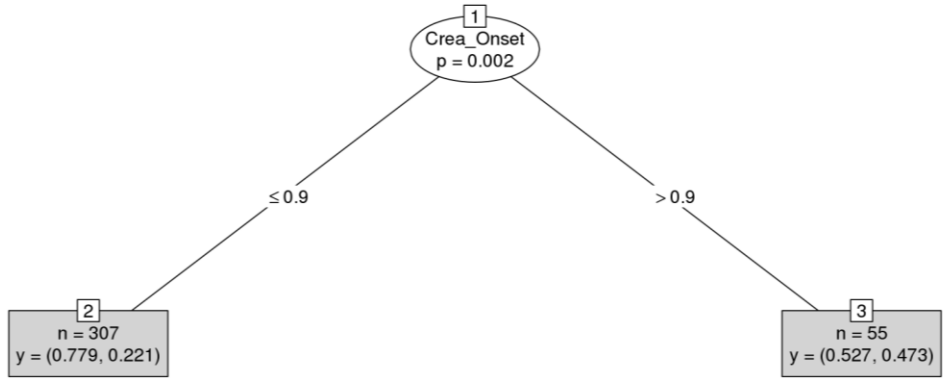
					Number of obs	= 347
					Wald chi2(5)	= 30.17
Penalized log likelihood = -156.07558					Prob > chi2	= 0.0000

icumorti	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
ESR_Onset	-.0102464	.0036469	-2.81	0.005	-.0173942	-.0030986
Na_Onset	-.049189	.0267759	-1.84	0.066	-.1016689	.0032909
Album_Onset	-.4160498	.206177	-2.02	0.044	-.8201494	-.0119503
Ferr_Onset	.000013	7.14e-06	1.83	0.068	-9.58e-07	.000027
Crea_Onset	.4856916	.1817647	2.67	0.008	.1294394	.8419438
_cons	6.885719	3.631956	1.90	0.058	-.2327844	14.00422

Tab. 1- Coefficients and p-values for the final regression model for severe course

Creatinine (OR 1.6; 95%CI 1.13 – 2.3), albumin (OR 0.65; 95%CI 0.44 – 0.98) and ESR levels (OR 0.98; 95%CI 0.98 – 0.99) independent predictors of severe course. CART analysis identified creatinine as the only predictive parameter, with 0.9 mg/dl as the cut-offs value distinguishing patients at low and high probability of severe outcome (Fig. 2). The prediction accuracy of the CART model was 0.6849; 95% CI: (0.5656, 0.7887), the Mcnemar's Test P-Value : 4.49e-06.

Fig. 2 - CART for severe outcome



3.2 Mortality

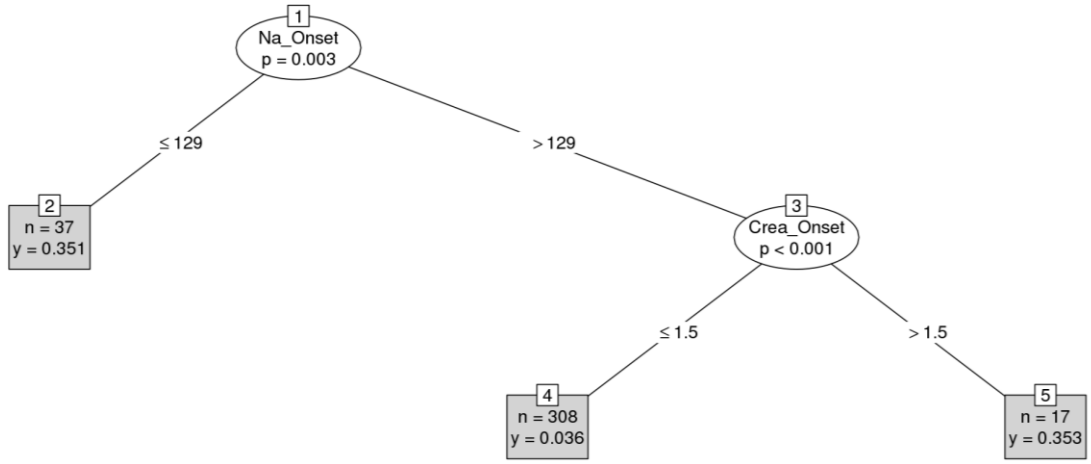
As shown in Tab.2, in the final regression model for mortality, serum sodium (OR 0.98; 96%CI 0.82-0.96) and creatinine values resulted the main predictors of death risk (OR 1.69; 96%CI 1.14-2.5).

		Number of obs	= 349
		Wald chi2(2)	= 15.37
Penalized log likelihood = -84.218312		Prob > chi2	= 0.0005

DeathYN	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Na_Onset	-.1115089	.0402107	-2.77	0.006	-.1903205	-.0326973
Crea_Onset	.5251773	.2002293	2.62	0.009	.132735	.9176195
_cons	12.13694	5.371665	2.26	0.024	1.608669	22.66521

Tab. 2 - Final regression model for mortality

CART analysis for mortality showed that subject most likely to die were those who presented with Na levels < 129 mg/dl and creatinine values > 1,3 mg/dl at onset, as



depicted in Fig. 3. The model accuracy was 0.9341, 95% CI : (0.862, 0.9754), with a Mcnemar's Test P-Value of 0.0412

Fig. 3 - CART for mortality

3.3 CNS disease

In the regression model for prediction of the development of CNS involvement age at MAS onset (OR 0.62; 95%CI 0.42 – 0.92) and fibrinogen levels (OR 0.99; 95%CI 0.99 – 0.99) were significantly associated with the outcome. Other variables in the model, albeit significantly associated, did not confer a substantial risk, as demonstrated by the odd ratios. The results are reported in Tab. 3

```
                                                Number of obs     =        349
                                                Wald chi2(8)      =      38.08
Penalized log likelihood = -148.60299           Prob > chi2       =     0.0000
```

CNS	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Age_at_onset	.4585325	.202439	2.27	0.024	.0617595	.8553056
Dis_dur_at_MAS_onset	.4710302	.197311	2.39	0.017	.0843078	.8577526
Age_at_MAS_onset	-.4687336	.1988192	-2.36	0.018	-.8584121	-.079055
AST_Onset	.0004154	.0001816	2.29	0.022	.0000595	.0007714
Fib_Onset	-.0020058	.0007048	-2.85	0.004	-.0033872	-.0006244
Ferr_Onset	.0000211	7.72e-06	2.73	0.006	5.96e-06	.0000362
BUN_Onset	.0123996	.0059299	2.09	0.037	.0007772	.0240219
Crea_Onset	.450411	.2512019	1.79	0.073	-.0419357	.9427577
_cons	-1.096789	.3544139	-3.09	0.002	-1.791427	-.4021505

Tab. 3 - Final regression model for CNS disease

Despite this, CART algorithm identified significant cut-offs of ferritin, creatinine and AST levels that are predictive of a high risk of CNS dysfunction. The model accuracy was 0.66. (Fig. 4).

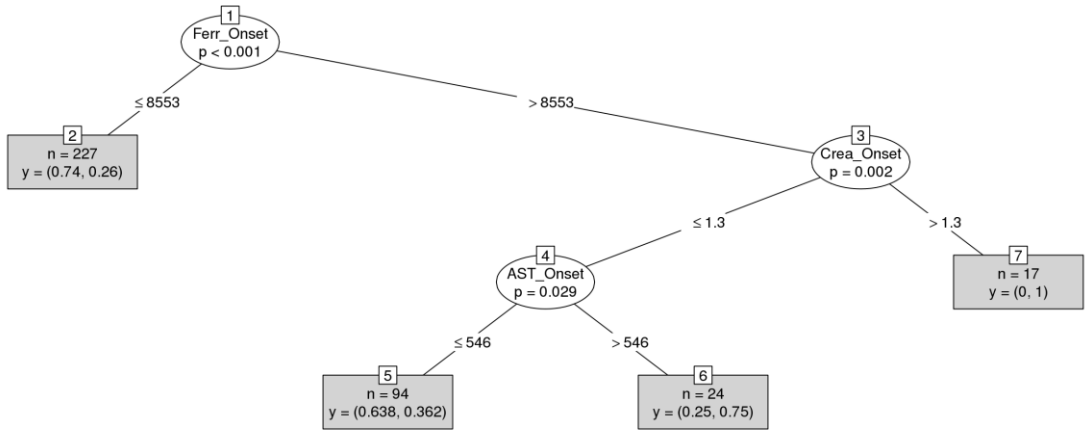


Fig. 4 – CART for CNS disease

4. Discussion

In this analysis, we identified the main important predictors of unfavorable outcomes and CNS dysfunction in sJIA-MAS patients. We took advantage of a causal inference framework for estimating unbiased results through the use of Directed Acyclic Graphs. This methodology is currently considered the state-of-the art for the analysis of observational data[50]. In complex disease like MAS, where outcome is affected by multiple factors and many analyzed variables presumably influences the others, correction of possible confounders is crucial for a correct estimation of risk. DAG-informed regression modelling allowed to identify some predictors of outcomes among variables routinely avaible in clinical practice at MAS onset. Of notice, our

[50] Staplin, Natalie, et al. "Use of causal diagrams to inform the design and interpretation of observational studies: an example from the Study of Heart and Renal Protection (SHARP)." *Clinical Journal of the American Society of Nephrology* 12.3 (2017): 546-552.

results differed from those of a previous analysis of the same cohort, where CNS involvement, heart, lung, or kidney failure, hemoglobin ≤ 7.9 g/dl, and age at onset of MAS > 11.5 years were independent predictors of severe course. Here, we aimed to predict MAS complications before they occurred; therefore, we did not include in the analysis the occurrence of organ failure, as the exact timing of occurrence with respect to disease onset was not asked for in the original survey. In the current work, nor hemoglobin nor age at MAS onset confirmed their role as predictors. This result could be due to the different regression technique, that was chosen to minimize small-sample bias, or due the use of DAGs to guide confounders adjustment. In our model, hyponatremia and high creatinine levels emerged as strong predictors of mortality, with cut-off values compatible with the occurrence of Acute Kidney Injury (AKI). Moreover, creatinine levels defined two risk distinct groups according to the CART model, with a threshold value of 0.9 mg/dl. These results suggest that kidney involvement could play a significant role in the occurrence of unfavorable outcomes in MAS. This finding is in line with the results of the latent subgroup analysis on the same sample presented in Chapter 2; moreover, it is consistent with evidences from the literature on HLH patients. The development of AKI has been reported as a risk factor for poor outcome in hemophagocytic syndromes[51]. Many pathogenetic mechanism have been invoked to explain such an involvement. AKI can occur as a consequence of hemodynamic dysfunction in the context of multiorgan failure (MOF) that can complicate the most severe MAS cases. However, kidney involvement is documented also in patients without

[51] Filippone, Edward J., and John L. Farber. "Hemophagocytic lymphohistiocytosis: an update for nephrologists." International urology and nephrology 48.8 (2016): 1291-1304.

MOF or shock. In most cases, AKI is attributed to acute tubular necrosis[52]. Thrombotic disease could be one of the possible causes of kidney damage. For example, thrombotic microangiopathy (TMA) is an increasingly recognized life-threatening condition that can superimpose to MAS, and is characterized by hemolytic anemia, decreased platelet counts, high LDH levels and AKI. However, also in absence of manifest thrombotic complications, MAS-related coagulopathy could play a role, as micro thrombosis of renal vessels has been reported in subject with disseminate intravascular coagulopathy (DIC) or DIC-like abnormalities like those seen in MAS[53]. In rare cases, renal injury could be also driven by histiocytic infiltration of the glomerular epithelium. Direct cytokinin damage, by TNFa and IL-6, has been proposed in absence of any hemodynamic or bioptic alteration[54]. Further researches are needed to investigate which of these mechanisms are more relevant and how kidney involvement should be addressed therapeutically in these patients. Regarding the occurrence of CNS dysfunction, hypofibrinogenemia resulted a significant risk factor. This suggest that coagulopathy and possibly hemorrhagic complication could be important determinants of neurologic complications in MAS. Another interesting finding is the inverse association of CNS with age at MAS onset, which suggests that younger patients are more susceptible of CNS involvement. Finally, although they didn't emerge as independent predictors for neurologic manifestations, ferritin, creatinine and AST values defined high risk-subgroups for CNS complications. In particular ferritin levels above

[52] Karras, Alexandre. "What nephrologists need to know about hemophagocytic syndrome." *Nature Reviews Nephrology* 5.6 (2009): 329.
[53] Sakamaki, Yusuke, et al. "Renal thrombotic microangiopathy in a patient with septic disseminated intravascular coagulation." BMC nephrology 14.1 (2013): 260.
[54] Wan, L., Bellomo, r., Di Giantomasso, D. & ronco, C. The pathogenesis of septic acute renal failure. Curr. Opin. Crit. Care 9, 496–502 (2003).

around 8500 ng/ml, creatinine values above 1.3 mg/dl and AST levels above around 550 U/l resulted the most predictive thresholds. These finding have to be interpreted with caution, considering the many limitations of the study. We analyzed retrospective data gathered though a web-based survey, and thus subjected to possible reporting bias and errors. Moreover, the effect of treatments on outcomes and possible interactions with laboratory variables were not assessed in the analysis. Despite these issues, we believe that the results could encourage further prognostic research in MAS and raise the awareness among physicians of the most important factors predicting the clinical course in MAS patients. The implications of the study are twofold; in a clinical practice context, the identification of the most predictive risk factors could improve patient stratification and assist decision-making promoting the efficacious selection of immunosuppressive and supportive treatments. In research, the classification of patients in risk group could help the balancing of study cohorts for a more accurate assessment of efficacy and safety questions.

5. Conclusions

In this study, we analyzed commonly avaible laboratory markers in sJIA-MAS patients and identified the most predictive factors and threshold values for the development of a complicate clinical course, CNS disease and mortality in a large multinational cohort. Further validation, possibly in prospective populations, is warranted for the development of adequate prognostic models in MAS.

References

- Buatois, V., Chatel, L., Cons, L., Lory, S., Richard, F., Guilhot, F., ... & Kosco-Vilbois, M. H. (2017). Use of a mouse model to identify a blood biomarker for IFNγ activity in pediatric secondary hemophagocytic lymphohistiocytosis. Translational Research, 180, 37-52.
  failure. Curr. Opin. Crit. Care 9, 496–502 (2003).
- Filippone, Edward J., and John L. Farber. "Hemophagocytic lymphohistiocytosis: an update for nephrologists." International urology and nephrology 48.8 (2016): 1291-1304.
- Girard-Guyonvarc'h, C., Palomo, J., Martin, P., Rodriguez, E., Troccaz, S., Palmer, G., & Gabay, C. (2018). Unopposed IL-18 signaling leads to severe TLR9-induced macrophage activation syndrome in mice. Blood, 131(13), 1430-1441.
- Heinze, Georg, et al. "Firth's bias reduced logistic regression." *R package version* 1 (2013).
- Karras, Alexandre. "What nephrologists need to know about hemophagocytic syndrome." *Nature Reviews Nephrology* 5.6 (2009): 329.
- Loh, Wei-Yin. "Classification and regression tree methods." *Wiley StatsRef: Statistics Reference Online* (2014).
- Mburu, Josephine W., et al. "Use of classification and regression tree (CART), to identify hemoglobin A1C (HbA1C) cut-off thresholds predictive of poor tuberculosis treatment outcomes and associated risk factors." *Journal of Clinical Tuberculosis and Other Mycobacterial Diseases* 11 (2018): 10-16.
- Minoia, Francesca, et al. "Dissecting the heterogeneity of macrophage activation syndrome complicating systemic juvenile idiopathic arthritis." The Journal of rheumatology 42.6 (2015): 994-1001.
- Ravelli, Angelo, et al. "Macrophage activation syndrome." Hematology/Oncology Clinics 29.5 (2015): 927-941.
- Sakamaki, Yusuke, et al. "Renal thrombotic microangiopathy in a patient with septic disseminated intravascular coagulation." BMC nephrology 14.1 (2013): 260.
- Sawhney S, Woo P, Murray KJ. Macrophage activation syndrome: a potentially fatal complication of rheumatic disorders. Arch Dis Child. 2001; 85:421–6
- Snowden, Jonathan M., et al. "The Curse of the Perinatal Epidemiologist: Inferring Causation Amidst Selection." *Current epidemiology reports* 5.4 (2018): 379-387.
- Staplin, Natalie, et al. "Use of causal diagrams to inform the design and interpretation of observational studies: an example from the Study of Heart and Renal Protection (SHARP)." *Clinical Journal of the American Society of Nephrology* 12.3 (2017): 546-552.
- T. Silfverberg, M. Arico, G. Janka, J.I. Henter, Risk factors for early death in children with haemophagocytic lymphohistiocytosis, Acta Paediatr. 101 (3) (2012)
- Tang, Y., & Luo, Z. (2018). Prognostic factors of early death in children with hemophagocytic lymphohistiocytosis. Cytokine, 110, 481-482. doi:10.1016/j.cyto.2017.09.023

- Vittinghoff, Eric, et al. *Regression methods in biostatistics: linear, logistic, survival, and repeated measures models*. Springer Science & Business Media, 2011.
- Wan, L., Bellomo, r., Di Giantomasso, D. & ronco, C. The pathogenesis of septic acute renal
- Williams, Thomas C., et al. "Directed acyclic graphs: a tool for causal studies in paediatrics." *Pediatric research* 84.4 (2018): 487.

Chapter 4 – Understanding causal pathways in SJIA-MAS: a causal inference approach to MAS prognosis

1. Introduction

Macrophage Activation Syndrome (MAS) is a potentially life-threatening hyper-inflammatory disorder residing in the group of Hemophagocytic Lymphohistiocytosis (HLH) that can complicate several systemic diseases, more frequently Systemic Juvenile Idiopathic Arthritis (SJIA). In some cases, MAS rapidly evolves into progressive multi-organ failure and eventually death [55].

A complex interrelationship of clinical, demographical, etiological and therapeutic factors determines in the development, progression, and outcome of this severe condition.

Despite the recently growing evidence of a key role of specific cytokines – namely IL-18 and INFg – in MAS development[56], most aspects of the pathogenetic pathways that govern the interindividual variation in severity, leading in a subgroup of patients to organ dysfunction and systemic widespread of the disease, remain obscure.

Currently, in the absence of evidence-based guidelines, the rationale for many therapeutic options in MAS is based on anecdotic data or inferred by studies on other forms within the HLH spectrum[57].

[55] Ravelli, A., et al. "Macrophage Activation Syndrome." Handbook of Systemic Autoimmune Diseases. Vol. 11. Elsevier, 2016. 85-106.
[56] Crayne, Courtney B., et al. "The Immunology of Macrophage Activation Syndrome." Frontiers in immunology 10 (2019).
[57] Davì, Sergio, et al. "Macrophage Activation Syndrome." Pediatric Rheumatology. Springer, Singapore, 2017. 275-292.

Understanding the causal contribution of different measurable variables to clinical outcomes is crucial for risk stratification, rational optimization of treatment sequences and life-saving adjuvant interventions, as well as for the development and assessment of novel therapeutic interventions.

Admission laboratory markers can provide useful information to predict the risk of a complicated clinical course in SJIA-MAS. However, prediction and causal inference are different concepts. The goal of prediction modelling is to develop useful tools to forecast the future occurrence of an outcome. Prediction studies do not necessarily include causal factors (targets of intervention), but may employ measure surrogates or biomarkers of causes. In the prediction context, causal inference is mainly used to incorporate other predictors was necessary to minimize confounding, so to obtain an unbiased estimation. Conversely, explanatory studies assess the role of causal factors which, if intervened upon through treatment, are likely to modify the outcomes[58] [59]. This means that models for causal inference are concerned with optimally deriving the likely change in an outcome due to potential change in specific variables, thus informing treatment choices[60]. However, both tasks must be considered in developing a useful prognostic model. As effectively pointed out by Gagliardi[61], to be adopted as risk factors, predictors must sit in the causal pathway of outcomes, so that their manipulation can cause a clinical improvement. Estimating causal

[58] Schooling, C. Mary, and Heidi E. Jones. "Clarifying questions about “risk factors”: predictors versus explanation." Emerging themes in epidemiology 15.1 (2018): 10.
[59] Arnold, Kellyn F., et al. "DAG-informed regression modelling, agent-based modelling and microsimulation modelling: a critical comparison of methods for causal inference." International journal of epidemiology 48.1 (2018): 243-253.
[60] Arnold, Kellyn F., et al. "Generalised linear models for prognosis and intervention: Theory, practice, and implications for machine learning." arXiv preprint arXiv:1906.01461 (2019).
[61] Gagliardi, Luigi. "Prediction and causal inference." Acta Paediatrica 98.12 (2009): 1890-1892.

effects from observational data, however, requires special methods, including causal mediation analysis and propensity scores[62].

In this study, we aim to explore the causal impact of the identified prognostic factors of outcomes in SJIA-MAS, by answering the following question: 'to what degree do the specific laboratory abnormalities increase the risk of developing unfavorable outcomes in MAS patients?'

For this task, we employed a structural model causal inference approach to unravel the causal relationships between demographical, laboratory, clinical and therapeutic variables in a the largest multinational SJIA-MAS cohort collected so far.

2. Methods

To estimate the causal contributions to the development of severe course, mortality and CNS disease, we evaluated the relationships between laboratory predictors at onset and outcomes within a structural causal inference framework[63]. This approach is based on counterfactual reasoning, i.e. it compares the potential outcomes between groups who exposed and not exposed to factors of interest[64].

Graphical models like directed acyclic graphs (DAGs) are used in this context to formalize the hypothesized causal relationships between variables, encoded as conditional independence

[62] E. Vittinghoff et al., Regression Methods in Biostatistics, Statistics for Biology and Health, DOI 10.1007/978-1-4614-1353-0 1

[63] Pearl J. Causality: Models, Reasoning and Inference. Cambridge, UK: Cambridge University Press; 2009.

[64] Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies. Journal of Educational Psychology. 1974; 66(5):688

relations in a probability distribution, the truthfulness of which can be assessed empirically in the dataset.

In DAGs, statistical dependencies between variables are represented as arrows between nodes. Variables influencing at least two others, including those thought to create or transmit spurious associations (confounders) should be added to the model; absence of an arrow between (observed or unmeasured) nodes indicates no causal effect[65]. Once the map of causal assumptions is constructed, mathematical rulesets known as d-separation criteria help to establish which variables should be controlled for when estimating the effect of one variable on another[66].

This formalism also allows distinguishing between the effect of the exposure on outcomes at a fixed level of the mediators – the direct effect – and the total effect which includes indirect effects via mediating variables[67]

We elaborated a separated DAGs of SJIA-MAS for each of the outcomes, accounting for avaible laboratory values at onset, clinical and demographic feature, geographic and specialty setting, and treatments. The minimal sufficient adjustment sets of variables that would allow the identification of an unconfounded effect of specific prognostic factors on outcomes were selected. We then queried the DAGs by applying the hill-climbing algorithm to the data and

[65] Morgan SL, Winship C, Counterfactuals and Causal Inference: Methods and Principles for Social Research . New York: Cambridge University Press, 2007

[66] Ferguson, Karl D., et al. "Evidence synthesis for constructing directed acyclic graphs (ESC-DAGs): a novel and systematic method for building directed acyclic graphs." International journal of epidemiology (2019).

[67] Leopold SJ, Watson JA, Jeeyapant A, Simpson JA, Phu NH, Hien TT, et al. (2019) Investigating causal pathways in severe falciparum malaria: A pooled retrospective analysis of clinical studies. PLoS Med 16(8): e1002858

comparing our theory-driven graphs to the obtained Bayesian network (BN) models, which are DAGs learned directly from data.

Based on emerged condition dependencies encoded in DAGs and BNs and the derived adjustment sets, we implemented a multivariable regression model to estimate the unbiased effect of candidate predictors on outcomes.

To evaluate possible non-linear and interdependent effects, we employed multivariate adaptive regression splines (MARS)[68] and partial dependence plots (PDP)[69].

Finally, to account for possible treatment confounders, marginal posterior probability for the treatment effect of cyclosporine and etoposide use on outcomes was estimated through augmented inverse probability weighting (AIPW) with double robust methods, which model both the outcome and the treatment probability so that only one of the two models must be correctly specified to consistently estimate the treatment effects.

R package MICE was used for data imputation. For DAGs and BNs implementation and we used DAGgity and the bnlearn packages through the Causal Analytics Toolkit (CAT) software (http://cox-associates.com/CAT.htm). Regression and mediation analyses was performed respectively in the MASS and medmod R packages, as implemented by Jamovi software. We used

[68] Friedman, Jerome H., and Charles B. Roosen. "An introduction to multivariate adaptive regression splines." (1995): 197-217.
[69] Greenwell, Brandon M. "pdp: an R Package for constructing partial dependence plots." The R Journal 9.1 (2017): 421-436.

Stata (version 13.0; StataCorp, LLC) for AIPW[70]. MARS models were implemented with the Salford Predictive Modeler (SPM) 8 software.

3. Results

The main goal of this analysis was to estimate unbiased causal effects on the occurrence of severe course, CNS disease and fatal outcome in MAS of the values of Hb (anemia), ferritin, ESR, AST (hepatic involvement), creatinine (AKI) at disease onset, as well as of age and SJIA duration at MAS diagnosis.

In the BNs models, we blacklisted[71] the arcs from outcomes to baseline laboratory variables and from these to the variables of treatments, continent, specialty and demographic features – as we assume that the former can't influence the latter.

Using d-separation rules and the obtained DAGs, we derived the minimal adjustments that allow for unbiased estimates of the causal effects.

Serum creatinine and Na levels (AKI) emerged as strong predictors of a fatal outcome, as described previously. Creatinine levels were the strongest risk factor for a severe course, followed by and LDH.

[70] Cattaneo, Matias D., David M. Drukker, and Ashley D. Holland. "Estimation of multivalued treatment effects under conditional independence." The Stata Journal 13.3 (2013): 407-450.

[71] The term blacklist refers to a set of relationships that, based on previous knowledge, can reasonably be excluded. This can regard the association between two variables or the direction of the causal relationship. To avoid the incorporation of incorrect links in BN, these relationships are ignored during structure learning.

In the following paragraphs the results from the graph models for each outcome, the adjusted regression models, as well as the non-linear models and partial dependent plots for the evaluation of non-linear relationships are detailed for the predictors of interest.

3.1 Evaluation of the causal effect of specific predictors on outcomes

3.1.1 Severe course

The DAG and estimated hill-climbing algorithm-based BN for severe course are depicted in Fig. 1 e 2.

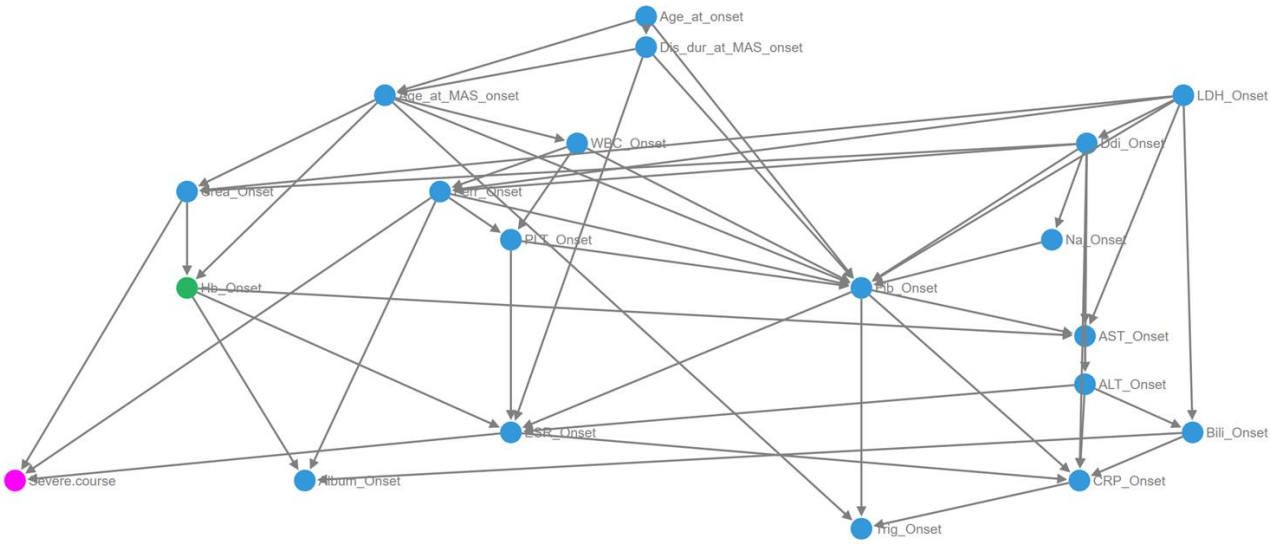


Fig. 1 – A priori DAG for severe course

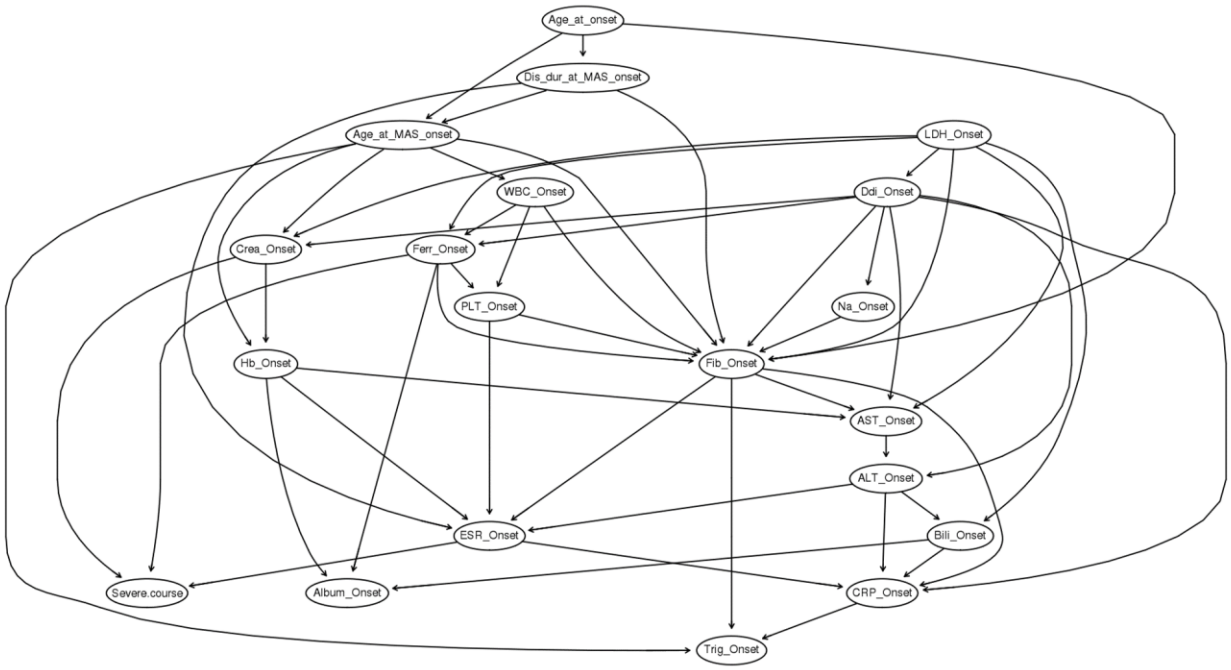


Fig. 2 – Bayesian Network for severe course

Age at sJIA onset

Age at sJIA onset resulted positively correlated with severe course after adjustment with AST, creatinine, d-dimer, disease duration at MAS onset, ferritin, fibrinogen, hemoglobin and platelet count, based on Spearman's rank correlation of 0.957 and p-value 0.00. However, the relationship resulted not significant in the adjusted regression model (Tab 1)

Predictor	Estimate	95% Confidence Interval		SE	Z	p	Odds ratio	95% Confidence Interval	
		Lower	Upper					Lower	Upper
Intercept	0.65219	-0.63168	1.93606	0.6550	0.9956	0.319	1.920	0.532	6.931
Age_at_onset	-0.03636	-0.09631	0.02359	0.0306	-1.1887	0.235	0.964	0.908	1.024
Dis_dur_at_MAS_onset	-0.11301	-0.19267	-0.03336	0.0406	-2.7807	0.005	0.893	0.825	0.967
Hb_Onset	-0.08524	-0.04428	0.21476	0.0661	1.2899	0.197	1.089	0.957	1.240
PLT_Onset	-3.31e-5	-0.00152	0.00158	7.91e-4	0.0418	0.967	1.000	0.998	1.002

Predictor	Estimate	95% Confidence Interval		SE	Z	p	Odds ratio	95% Confidence Interval	
		Lower	Upper					Lower	Upper
Fib_Onset	0.00193	4.18e-4	0.00345	7.73e-4	2.501	0.012	1.002	1.000	1.003
Ferr_Onset	-1.60e−5	-2.98e−5	-2.17e−6	7.04e-6	-2.270	0.023	1.000	1.000	1.000
Crea_Onset	-0.44222	-0.82539	-0.05904	0.1955	-2.262	0.024	0.643	0.438	0.943

Tab. 1- Regression analysis for the effect of Age at sJIA onset on severe course

Partial dependence plots (PDPs) reveal an increase of risk for patients with SJIA onset at 2 years of age, and again above 12 years of age (Fig 3).

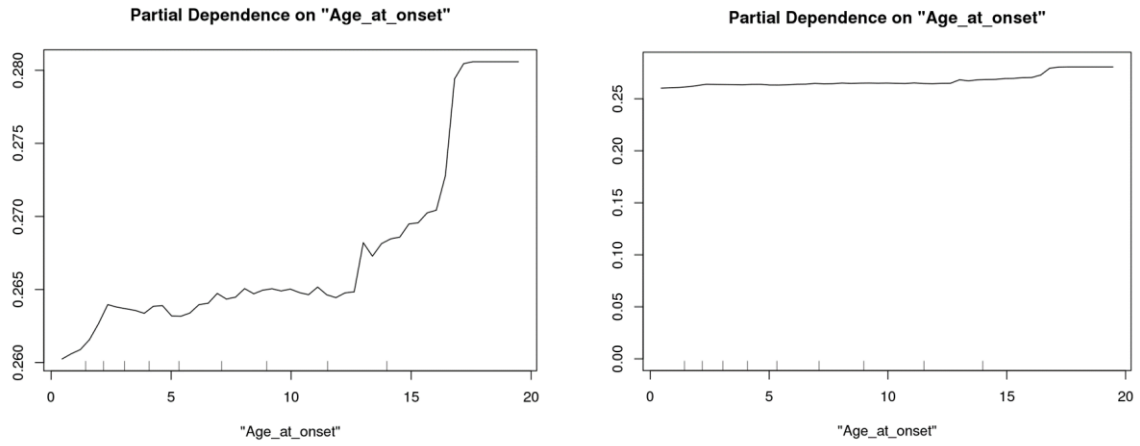


Fig. 3 -Partial dependence plots for age at sJIA onset. The plot shows the added-variable plots (depicting the relationship between the response variable and one of the predictors in the regression model, after controlling for the presence of the other predictors) for the parameters of interest.

A similar pattern emerged from the non-linear MARS plot (Fig. 4)

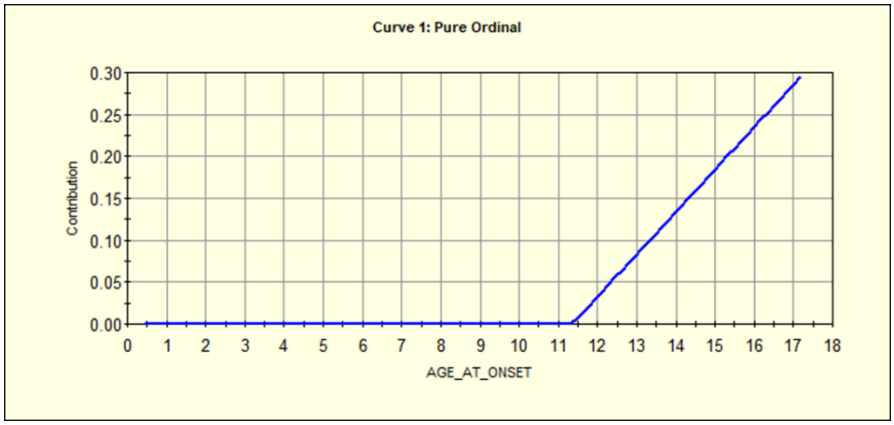


Fig. 4 – MARS model for Age at sJIA onset

Age at MAS onset

With an adjustment set of AST, creatinine, d-dimer, disease duration at MAS onset, ferritin, fibrinogen, Hb and PLT, Age at MAS onset was positively correlated to the risk of severe course

(based on Spearman's rank correlation of 0.988 and p-value 0.000), as shown in PDPs (Fig. 5); however, no significant relationship was observed through regression modelling (Tab. 2)

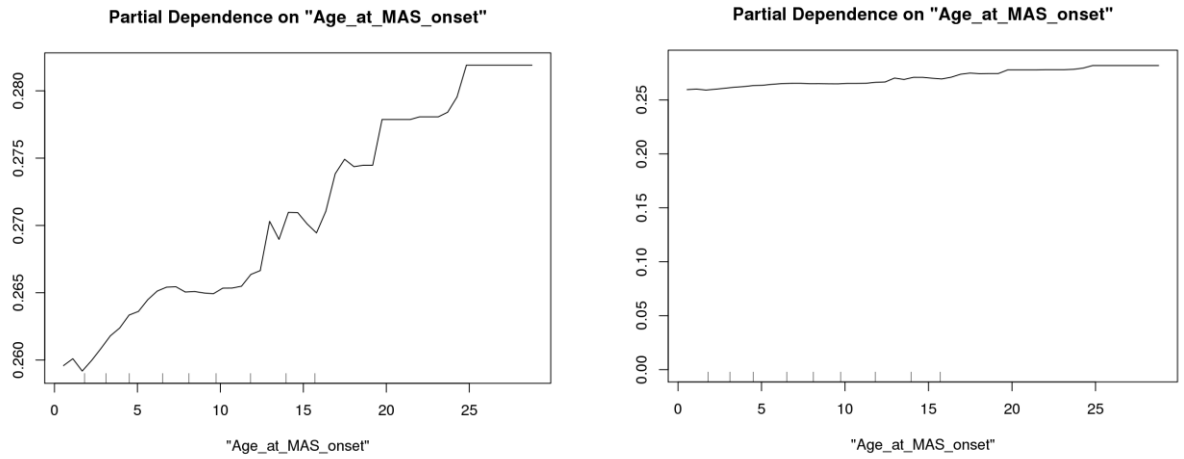


Fig. 5 – PDPs for age at MAS onset

		95% Confidence Interval						95% Confidence Interval	
Predictor	**Estimate**	**Lower**	**Upper**	**SE**	**Z**	**p**	**Odds ratio**	**Lower**	**Upper**
Intercept	0.68617	-0.60763	1.97998	0.6601	1.0395	0.299	1.986	0.545	7.243

Predictor	Estimate	95% Confidence Interval		SE	Z	p	Odds ratio	95% Confidence Interval	
		Lower	Upper					Lower	Upper
Age_at_onset	-0.29030	-0.67485	0.09426	0.1962	-1.4796	0.139	0.748	0.509	1.099
Dis_dur_at_MAS_onset	-0.35703	-0.73343	0.01938	0.1920	-1.8591	0.063	0.700	0.480	1.020
Hb_Onset	0.09442	-0.03737	0.22622	0.0672	1.4042	0.160	1.099	0.963	1.254
PLT_Onset	-7.30e−5	-0.00163	0.00148	7.93e-4	-0.0920	0.927	1.000	0.998	1.001
Fib_Onset	0.00200	4.04e-4	0.00360	8.14e-4	2.457	0.014	1.002	1.000	1.004

Predictor	Estimate	95% Confidence Interval		SE	Z	p	Odds ratio	95% Confidence Interval	
		Lower	Upper					Lower	Upper
Ferr_Onset	-1.48e−5	-2.91e−5	-4.38e−7	7.31e-6	-2.0199	0.043	1.000	1.000	1.000
Crea_Onset	-0.44667	-0.84479	-0.04855	0.2031	-2.1990	0.028	0.640	0.430	0.953
Age_at_MAS_onset	0.24686	-0.13001	0.62372	0.1923	1.2838	0.199	1.280	0.878	1.866
Ddi_Onset	5.96e-7	-3.83e−5	3.94e-5	1.98e-5	0.0301	0.976	1.000	1.000	1.000
AST_Onset	-1.82e−4	-4.43e−4	7.95e-5	1.33e-4	-1.3636	0.173	1.000	1.000	1.000

Tab. 5 – regression model for age at MAS onset

Dispite a similar trend showing an increase of risk above around 7 years of age, MARS model confirmed the absence of significant assosiation (Fig 6)

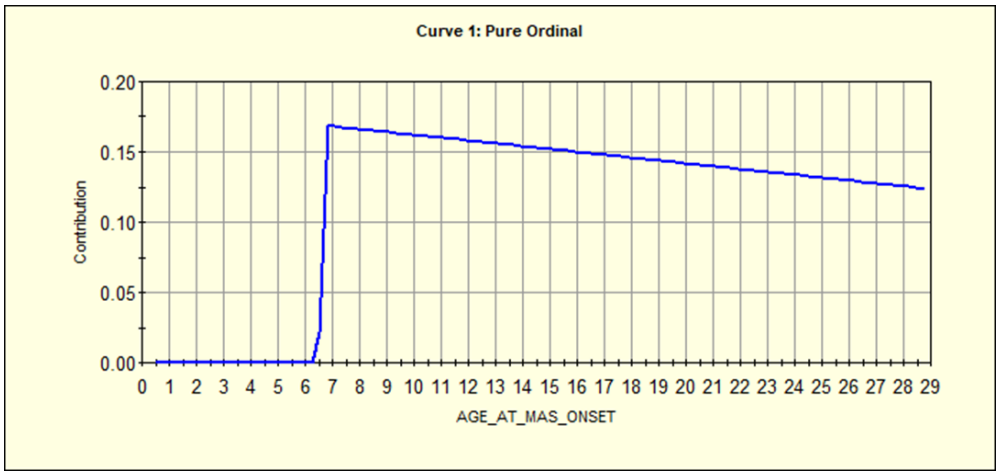


Fig. 6 – MARS model for age at MAS onset

Disease duration at MAS onset

Using an adjustment set of age at MAS onset, d-dimer, ESR, ferritin, Hb, LDH, SJIA duration at MAS onset appeared positively correlated based on Spearman's test (r 0.986 and p-value 0.000). PDPs showing an icrease of risk above around 10 years of disease duration are decipted in Fig. 7. However, the association resulted not significant in the regression analysis (p 0.089), as reported in Tab. 6. MARS model showed a not significant progressively rising of risk above around 3 years of disease duration.

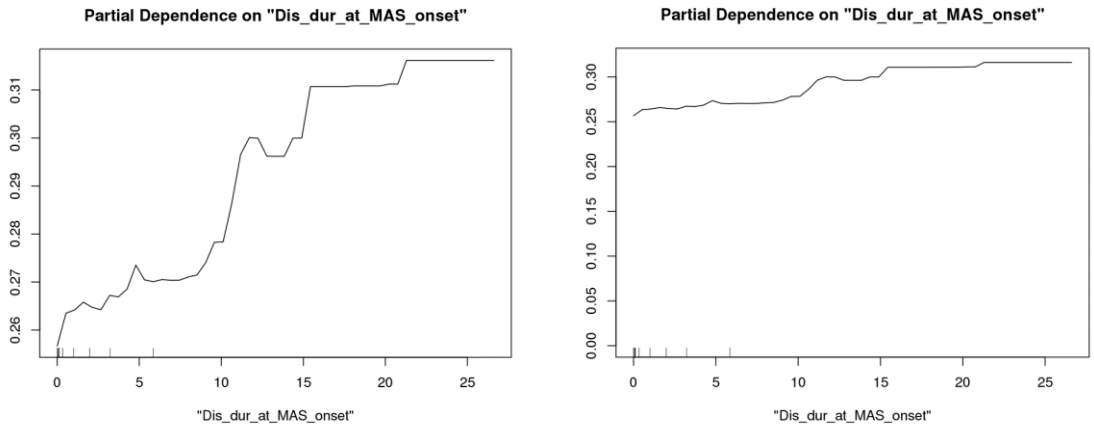


Fig. 7 – PDPs for disease duration at MAS onset

		95% Confidence Interval						95% Confidence Interval	
Predictor	**Estimate**	**Lower**	**Upper**	**SE**	**Z**	**p**	**Odds ratio**	**Lower**	**Upper**
Intercept	1.38187	0.81426	1.94949	0.28960	4.77	< .001	3.982	2.258	7.025
Dis_dur_at_MAS_onset	-0.06137	-0.13142	0.00867	0.03574	-1.72	0.086	0.940	0.877	1.009

Predictor	Estimate	95% Confidence Interval		SE	Z	p	Odds ratio	95% Confidence Interval	
		Lower	Upper					Lower	Upper
Ferr_Onset	-1.84e−5	-3.20e−5	-4.84e−6	6.94e-6	-2.66	0.008	1.000	1.000	1.000
Crea_Onset	-0.52310	-0.89712	-0.14908	0.19083	-2.74	0.006	0.593	0.408	0.862
ESR_Onset	0.00839	0.00109	0.01569	0.00372	2.25	0.024	1.008	1.001	1.016

Tab. 6 – Regression model for disease duration at MAS onset

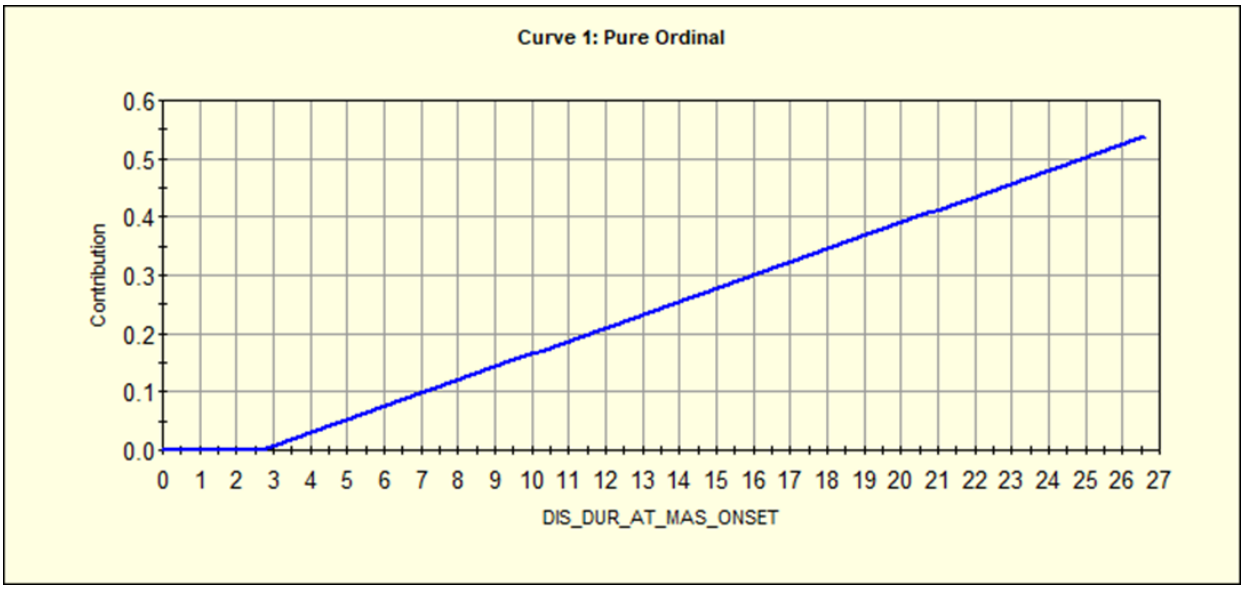


Fig. 8 – MARS model for disease duration at MAS onset

Hemoglobin

To assess Hb (anemia) direct effect on the risk of severe course, the DAG suggested a minimal adjustment set containing the following variables: age at MAS onset, disease duration at MAS onset, creatinine, d-dimer, ESR, fibrinogen, LDH, PLT.

No significant association was detected, as shown by the partial dependence plots (Fig 9). Results from the logistic model are reported in Table 7.

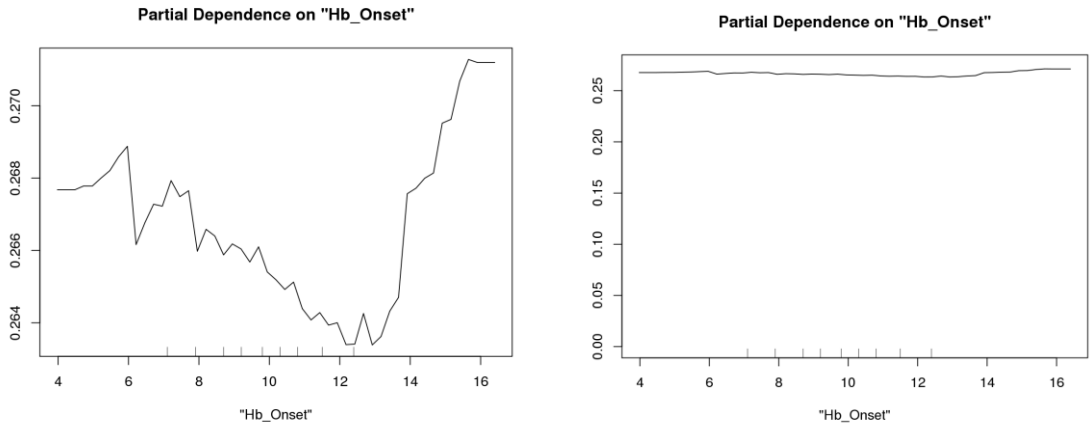


Fig. 9 – PDPs for Hb

Predictor	Estimate	95% Confidence Interval		SE	Z	p	Odds ratio
		Lower	Upper				
Intercept	0.03825	-1.38945	1.46595	0.72843	0.0525	0.958	1.039
Age_at_MAS_onset	-0.03144	-0.09149	0.02862	0.03064	-1.0260	0.305	0.969
Dis_dur_at_MAS_onset	-0.06572	-0.14839	0.01694	0.04218	-1.5583	0.119	0.936
Crea_Onset	-0.47039	-0.87668	-0.06409	0.20730	-2.2691	0.023	0.625

Predictor	Estimate	95% Confidence Interval		SE	Z	p	Odds ratio
		Lower	Upper				
Ddi_Onset	-9.70e−6	-4.79e−5	2.85e-5	1.95e-5	-0.4981	0.618	1.000
ESR_Onset	0.00652	-0.00210	0.01513	0.00439	1.4830	0.138	1.007
Fib_Onset	0.00163	-5.96e−6	0.00327	8.37e-4	1.9528	0.051	1.002
PLT_Onset	-1.61e−4	-0.00172	0.00140	7.98e-4	-0.2014	0.840	1.000
LDH_Onset	-2.46e−5	-1.07e−4	5.77e-5	4.20e-5	-0.5855	0.558	1.000
Hb_Onset	0.11423	-0.02169	0.25016	0.06935	1.6472	0.100	1.121

Predictor	Estimate	95% Confidence Interval		SE	Z	p	Odds ratio
		Lower	Upper				

Note. Estimates represent the log odds of "severe course = 0" vs. "severe course = 1"

Tab. 7 – Regression model fo Hb

MARS model confimed the non-significance of the relationships between Hb and severe course, while showing a non-linear tendency with a progressive descrease of severe course risk for Hb values above around 9.5 g/dl (Fig. 10)

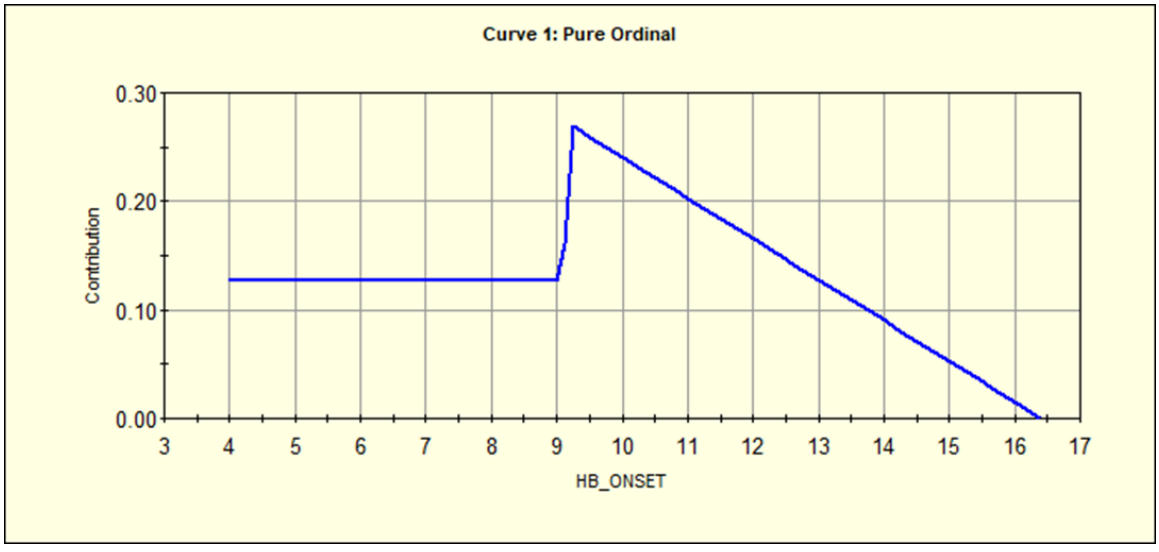


Fig. 10 – MARS model for Hb

Ferritin

Adjusting for creatinine and ESR values, a significant direct effect of Ferritin levels on the occurrence of a severe course was observed; however, the odd ratio did not actually suggest an increased risk depending from ferritin levels (Table 8).

		95% Confidence Interval						95% Confidence Interval	
Predictor	**Estimate**	**Lower**	**Upper**	**SE**	**Z**	**p**	**Odds ratio**	**Lower**	**Upper**
Intercept	1.18407	0.67270	1.69543	0.26091	4.54	< .001	3.268	1.960	5.449
ESR_Onset	0.00964	0.00252	0.01684	0.00364	2.65	0.008	1.010	1.003	1.017
Ferr_Onset	-1.72e−5	-3.08e−5	-3.73e−6	6.90e-6	-2.50	0.012	1.000	1.000	1.000
Crea_Onset	-0.54017	-0.91193	-0.16843	0.18968	-2.85	0.004	0.583	0.402	0.845

Tab. 8 – Regression model for ferritin

However, a significantly positive relationship was revealed by Spearman's rank correlation (r 0.827, p-0.000). PDPs reveal that the risk increases with ferritin levels until it reaches a plateau at around 20000 ng/dl, to increase again above levels of around 70000 ng/dl. This pattern was confirmed when accounting for the non-linear relationship through MARS regression (Fig 11)

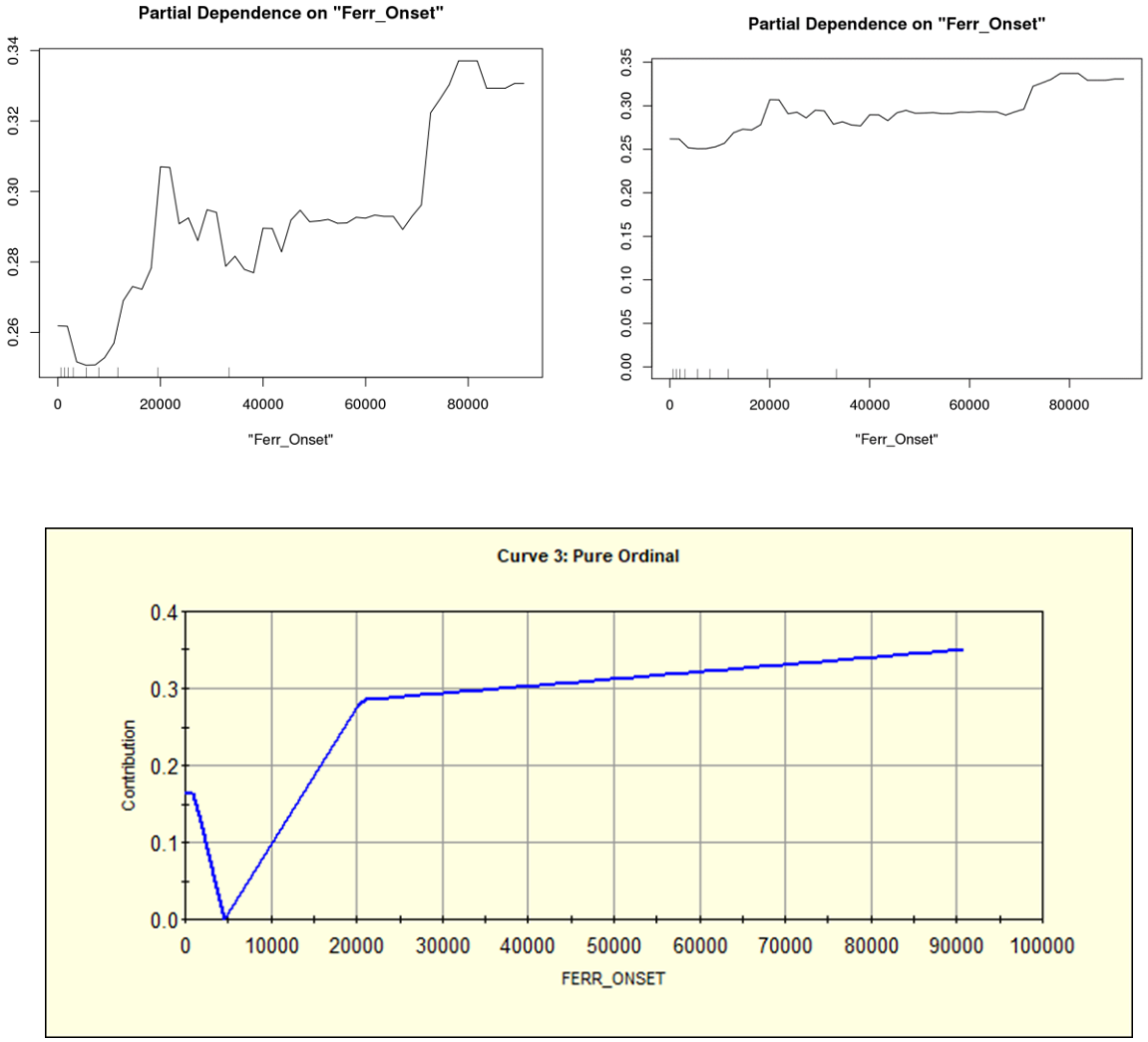


Fig. 11 – PDPs and MARS model for Ferritin

ESR

With an adjustment set of age at MAS onset, creatinine, d-dimer, disease duration at MAS onset, fibrinogen, LDH and PLT, the direct causal effect of ESR on severe course was not-significant (Table 9). PDPs, shown in Fig.12, and MARS regression (not shown) confirmed the absence of a significant association.

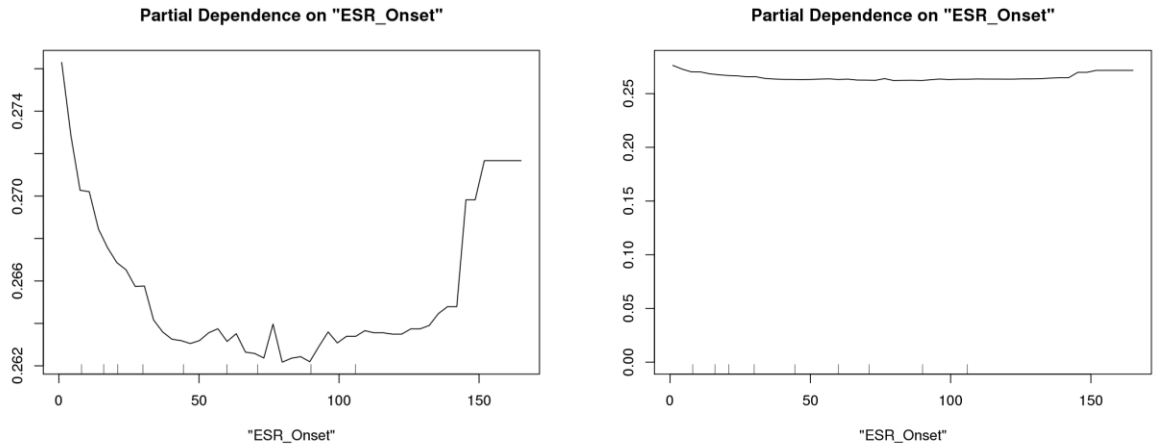


Fig. 12 – PDPs model for ESR

Tab. 9 – Below, the regression model for ESR

Predictor	Estimate	95% Confidence Interval		SE	Z	p	Odds ratio	95% Confidence Interval	
		Lower	Upper					Lower	Upper
Intercept	0.98197	0.21709	1.74685	0.39025	2.516	0.012	2.670	1.242	5.736
Age_at_onset	-0.04048	-0.09623	0.01526	0.02844	-1.423	0.155	0.960	0.908	1.015
Dis_dur_at_MAS_onset	-0.09160	-0.16896	-0.01425	0.03947	-2.321	0.020	0.912	0.845	0.986
LDH_Onset	-4.33e−5	-1.23e−4	3.66e-5	4.08e-5	-1.062	0.288	1.000	1.000	1.000
Ddi_Onset	-1.84e−5	-5.47e−5	1.78e-5	1.85e-5	-0.997	0.319	1.000	1.000	1.000
Fib_Onset	0.00189	2.60e-4	0.00351	8.29e-4	2.274	0.023	1.002	1.000	1.004

Predictor	Estimate	95% Confidence Interval		SE	Z	p	Odds ratio	95% Confidence Interval	
		Lower	Upper					Lower	Upper
PLT_Onset	1.11e-4	-0.00145	0.00167	7.98e-4	0.139	0.889	1.000	0.999	1.002
ESR_Onset	0.00324	-0.00484	0.01133	0.00412	0.787	0.431	1.003	0.995	1.011

AST

No adjustment set resulted necessary to estimate the causal effect of hepatic disease (AST values) on severe course. As evident by odd ratio intervals, no significant relationship was demonstrated.

Predictor	Estimate	95% Confidence Interval		SE	Z	p	Odds ratio	95% Confidence Interval	
		Lower	Upper					Lower	Upper
Intercept	1.13	0.871	1.39	0.133	8.50	< .001	3.100	2.388	4.024
AST_Onset	-2.64e−4	-5.04e−4	-2.39e−5	1.22e-4	-2.16	0.031	1.000	0.999	1.000

Tab 10 – regression model for AST

Albumin

No adjustment set was required. Results from the regression model are shown in Tab. 11

		95% Confidence Interval						95% Confidence Interval	
Predictor	**Estimate**	**Lower**	**Upper**	**SE**	**Z**	**p**	**Odds ratio**	**Lower**	**Upper**
Intercept	-0.730	-1.932	0.471	0.613	-1.19	0.234	0.482	0.145	1.60
Album_Onset	0.567	0.178	0.955	0.198	2.86	0.004	1.762	1.195	2.60

Tab 11 – regression model for albumine

Fibrinogen

No adjustment set was required for the direct effect estimation. A significant, inverse relatioship between fibrinogen levels and the risk of severe course was detected (Tab.12)

		95% Confidence Interval						95% Confidence Interval	
Predictor	**Estimate**	**Lower**	**Upper**	**SE**	**Z**	**p**	**Odds ratio**	**Lower**	**Upper**
Intercept	-1.067	-1.315	-0.818	0.127	-8.40	< .001	0.344	0.268	0.441

		95% Confidence Interval						95% Confidence Interval	
Predictor	**Estimate**	**Lower**	**Upper**	**SE**	**Z**	**p**	**Odds ratio**	**Lower**	**Upper**
Fib_Onset	-0.443	-0.722	-0.164	0.142	-3.12	0.002	0.642	0.486	0.848

Tab 12 – regression model for fibrinogen

LDH

The direct effect could be estimated without adjustments. A significant impact of LDH levels on severe course was obeserved (Tab.13)

		95% Confidence Interval						95% Confidence Interval	
Predictor	**Estimate**	**Lower**	**Upper**	**SE**	**Z**	**p**	**Odds ratio**	**Lower**	**Upper**
Intercept	-1.035	-1.2768	-0.794	0.123	-8.40	< .001	0.355	0.279	0.452
LDH_Onset	0.299	0.0631	0.534	0.120	2.48	0.013	1.348	1.065	1.706

		95% Confidence Interval						95% Confidence Interval	
Predictor	**Estimate**	**Lower**	**Upper**	**SE**	**Z**	**p**	**Odds ratio**	**Lower**	**Upper**

Tab 13 – regression model for LDH

Creatinine

No necessary adjustment set emerged from the DAG. Regression modelling confimed the positive relationship with the outcome. MARS model revealed an increase of risk for serum creatinine > 0.7 mg/dl (Fig. 12)

		95% Confidence Interval						95% Confidence Interval	
Predictor	**Estimate**	**Lower**	**Upper**	**SE**	**Z**	**p**	**Odds ratio**	**Lower**	**Upper**
Intercept	-1.484	-1.853	-1.11	0.188	-7.88	< .001	0.227	0.157	0.328
Crea_Onset	0.636	0.256	1.02	0.194	3.28	0.001	1.889	1.292	2.762

Tab 13 – regression model for creatinine

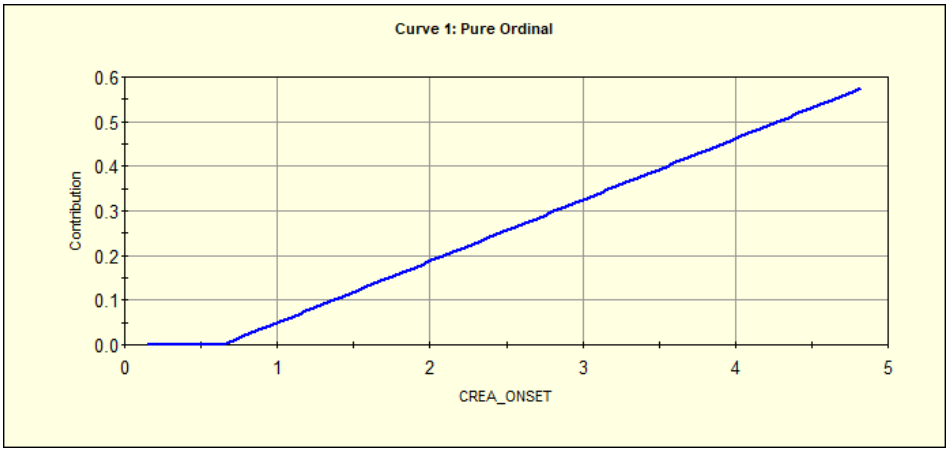


Fig. 12 – MARS model for creatinine

The relationship between the distribution of selected parameters the probability of outcome is summarized in the plots in Fig.13

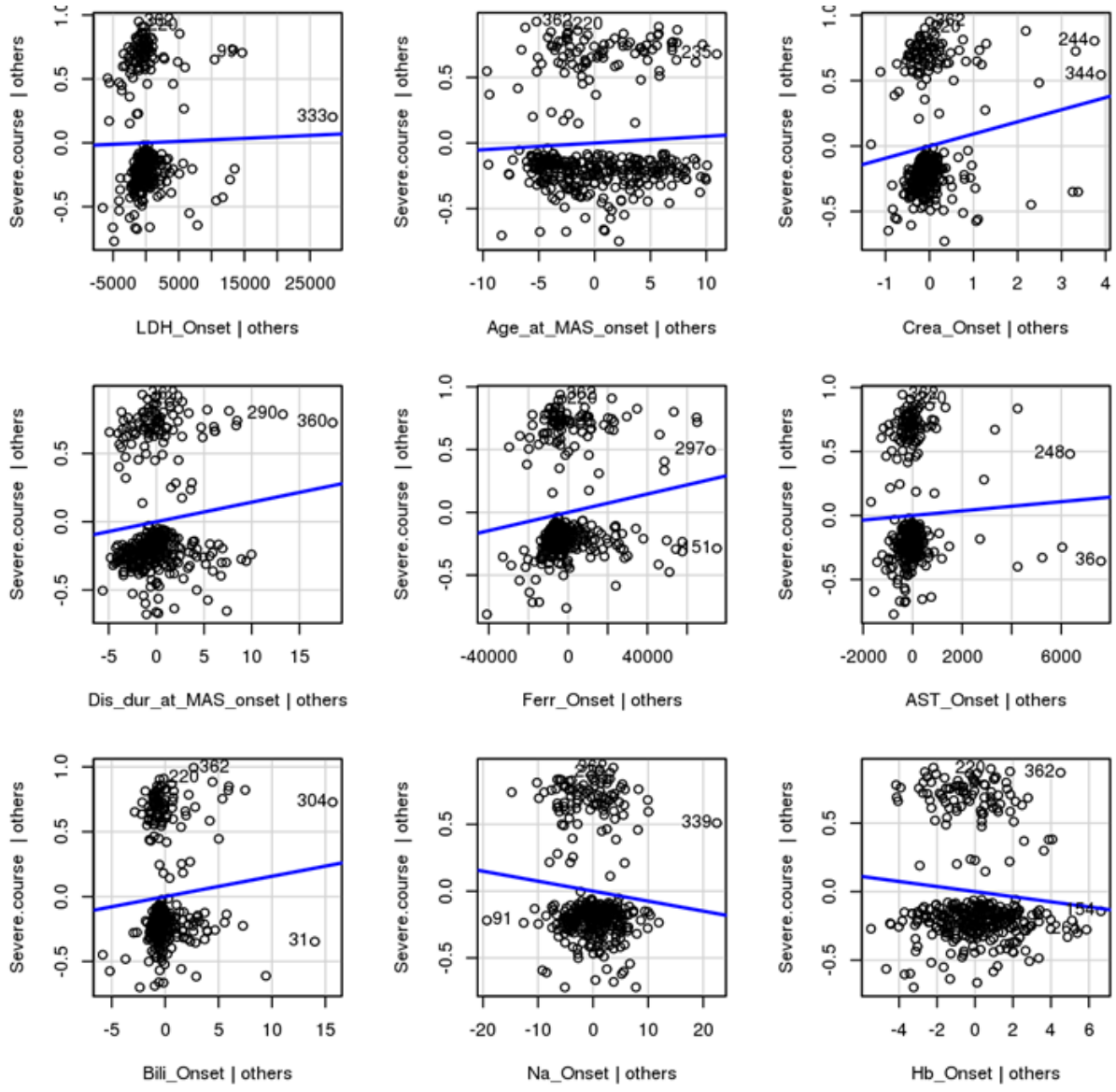


Fig- 13 - Added Variable Plots show the relationship between the response variable and one of the predictors in the regression model, after controlling for the presence of the other predictors

3.1.2 Death

The DAG and BN for death are represented in Fig 14.

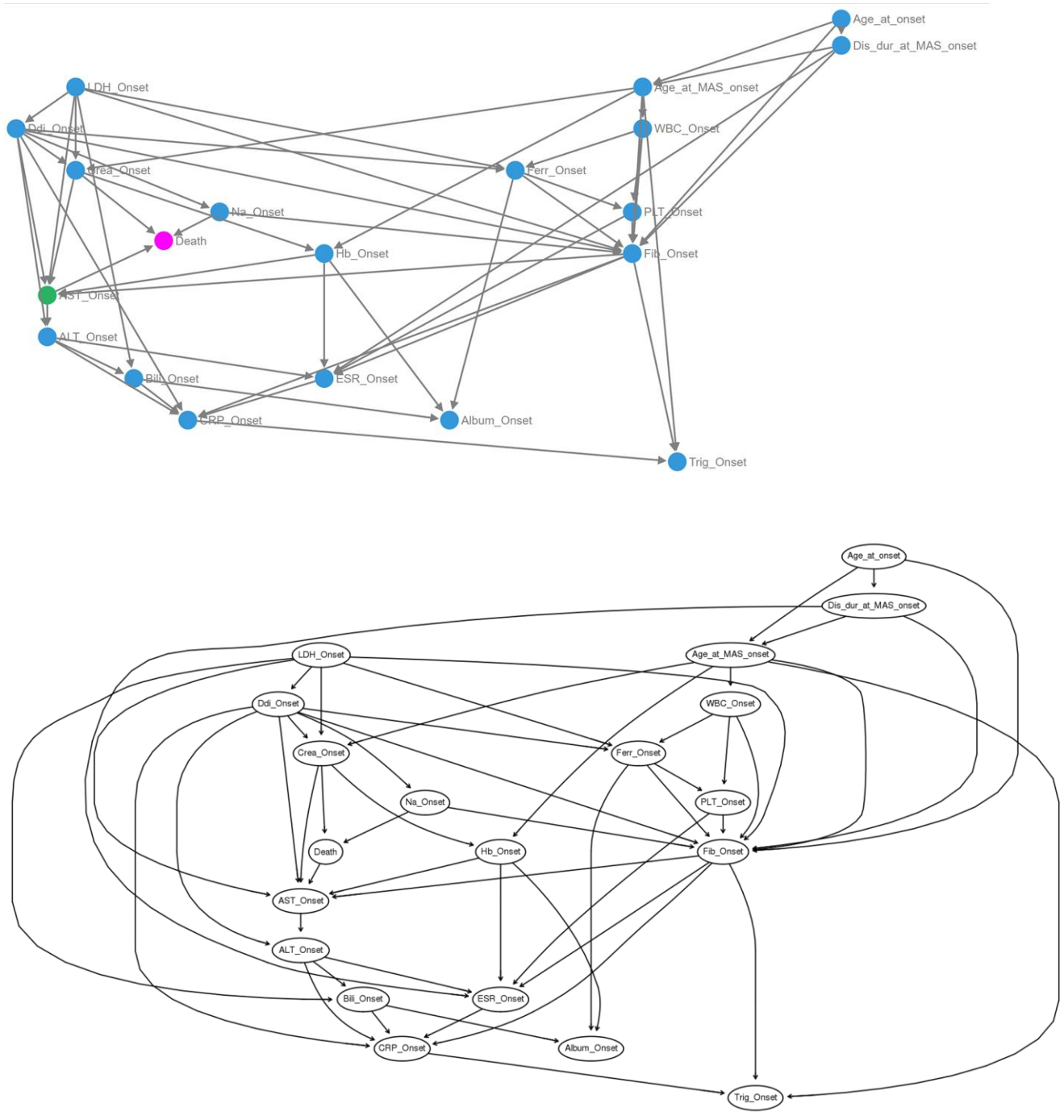


Fig. 14 – A priori and data-drive DAGs for mortality

AST

To assess the effect of hepatic involvement on death risk, according to these DAGs, we selected an adjustment set including age at MAS onset, creatinine, d-dimer, fibrinogen, LDH. Despite not being an independent predictor of death in the linear regression model (Tab. 14), AST levels positively correlated with mortality based on Spearman's test, with r 0.894 and p-value 0.000. The non-linear association is illustrated in PDPs, which show a rapid increase in death rate for AST values > 4000 U/L. MARS model confirmed the non-linear nature of the association, with an increase in death risk above values of around 2500 U/L (Fig. 15)

		95% Confidence Interval						95% Confidence Interval	
Predictor	**Estimate**	**Lower**	**Upper**	**SE**	**Z**	**p**	**Odds ratio**	**Lower**	**Upper**
Intercept	-2.01055	-3.09713	-0.924	0.55439	-3.627	< .001	0.134	0.0452	0.397
AST_Onset	1.77e-4	-1.17e−4	4.71e-4	1.50e-4	1.182	0.237	1.000	0.9999	1.000

Predictor	Estimate	95% Confidence Interval		SE	Z	p	Odds ratio	95% Confidence Interval	
		Lower	Upper					Lower	Upper
Ddi_Onset	4.01e-5	-1.56e−5	9.59e-5	2.84e-5	1.412	0.158	1.000	1.0000	1.000
Fib_Onset	-0.00300	-0.00565	-3.52e−4	0.00135	-2.221	0.026	0.997	0.9944	1.000
LDH_Onset	-1.30e−4	-3.11e−4	5.09e-5	9.23e-5	-1.408	0.159	1.000	0.9997	1.000
Age_at_MAS_onset	0.03655	-0.03724	0.110	0.03765	0.971	0.332	1.037	0.9634	1.117

Tab. 14 – regression model for AST

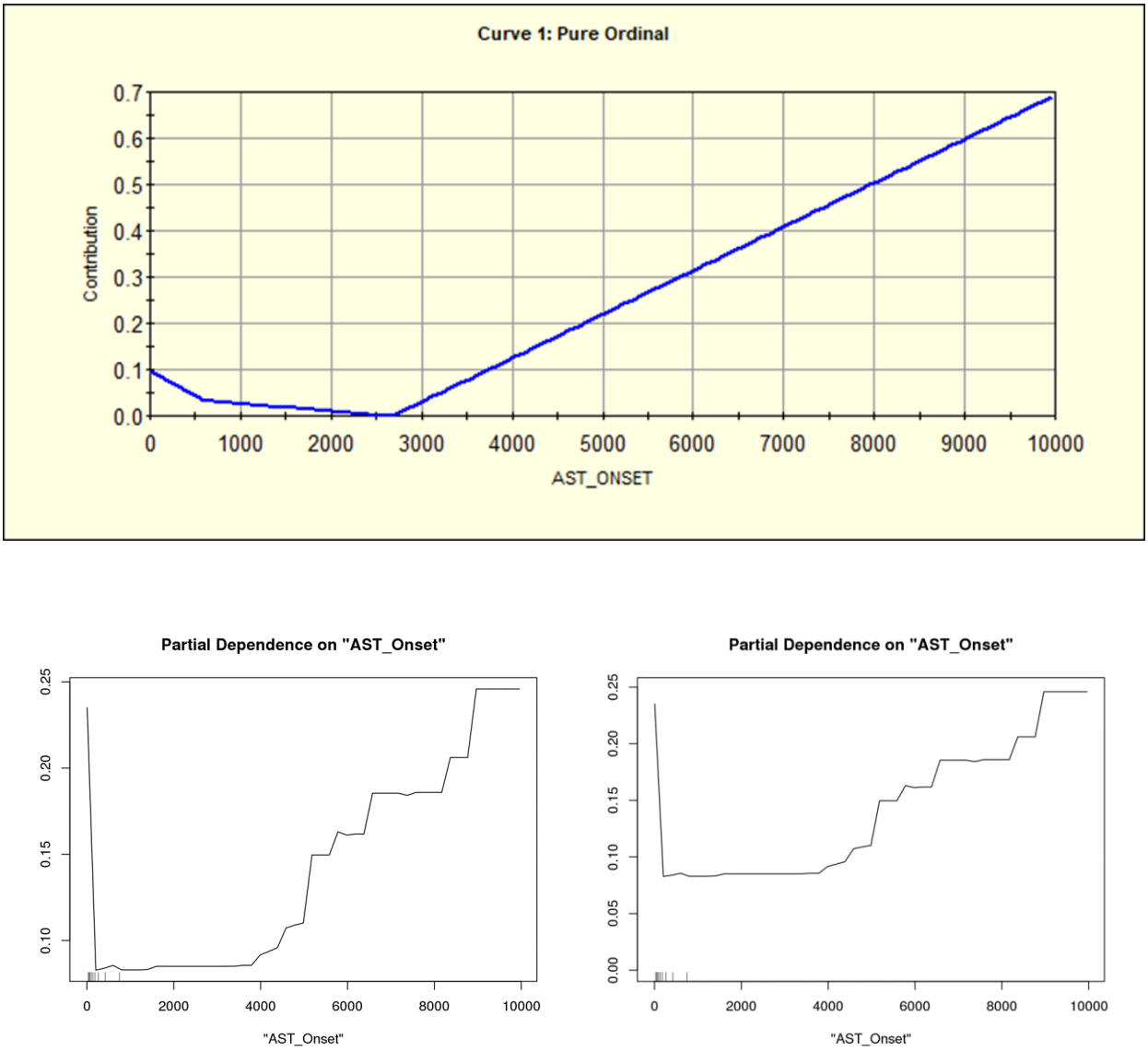


Fig. 15 – MARS model and PDPs for AST

Hemoglobin

An adjustment set of AST, age at MAS onset, creatinine, d-dimer, fibrinogen, LDH was considered. No significant association emerged, as revealed by PDPs (Fig. 16), regression modelling (Tab. 15) and MARS (not shown).

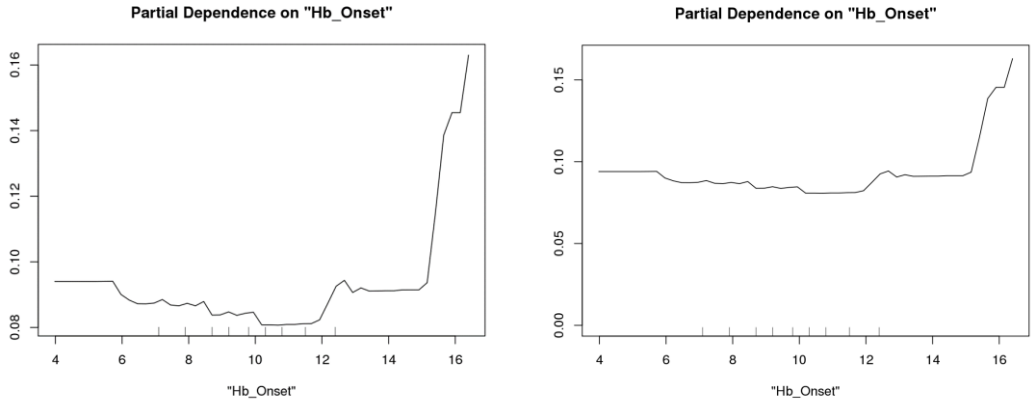


Fig. 16 – PDPs for Hb

Model Coefficients

Predictor	Estimate	95% Confidence Interval		SE	Z	p	Odds ratio	95% Confidence Interval	
		Lower	Upper					Lower	Upper
Intercept	-1.73356	-3.61880	0.152	0.96188	-1.802	0.072	0.177	0.0268	1.164
Fib_Onset	-0.00277	-0.00543	-1.03e−4	0.00136	-2.036	0.042	0.997	0.9946	1.000

Model Coefficients

Predictor	Estimate	95% Confidence Interval		SE	Z	p	Odds ratio	95% Confidence Interval	
		Lower	Upper					Lower	Upper
LDH_Onset	-1.65e−4	-3.56e−4	2.65e-5	9.75e-5	-1.688	0.091	1.000	0.9996	1.000
Age_at_MAS_onset	0.03063	-0.05631	0.118	0.04436	0.691	0.490	1.031	0.9452	1.125
AST_Onset	2.58e-4	-4.71e−5	5.64e-4	1.56e-4	1.658	0.097	1.000	1.0000	1.001
Ddi_Onset	2.70e-5	-3.03e−5	8.42e-5	2.92e-5	0.923	0.356	1.000	1.0000	1.000
Crea_Onset	0.59830	0.12592	1.071	0.24101	2.482	0.013	1.819	1.1342	2.917

Model Coefficients

Predictor	Estimate	95% Confidence Interval		SE	Z	p	Odds ratio	95% Confidence Interval	
		Lower	Upper					Lower	Upper
Hb_Onset	-0.06758	-0.26607	0.131	0.101247	-0.667	0.505	0.935	0.7664	1.140

Tab. 15 – regression model for Hb

Creatinine

When adjusted for Na levels, creatinine showed a positive correlation with mortality (Spearman's rank correlation 0.781, p-value 0.000). PDPs illustrate the increase in death risk observed for creatine levels above around 1.5 mg/dl (Fig. 17). Regression confirmed the relationship between creatinine and death risk (Tab. 16)

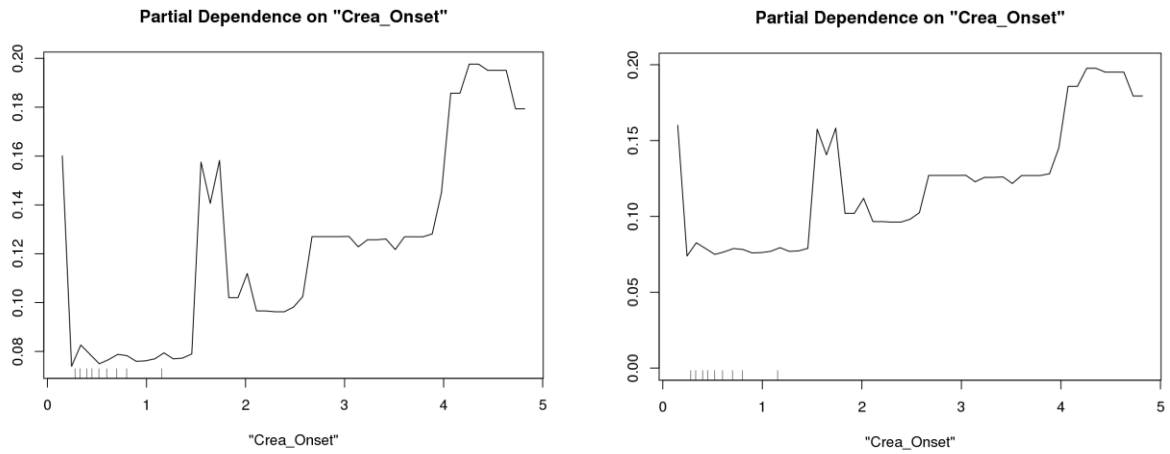


Fig. 17 – PDPs for creatinine

		95% Confidence Interval						95% Confidence Interval	
Predictor	**Estimate**	**Lower**	**Upper**	**SE**	**Z**	**p**	**Odds ratio**	**Lower**	**Upper**
Intercept	12.358	1.647	23.0691	5.4648	2.26	0.024	232880.229	5.194	1.04e+10
Crea_Onset	0.519	0.1066	0.9316	0.2105	2.46	0.014	1.680	1.112	2.539

Predictor	Estimate	95% Confidence Interval		SE	Z	p	Odds ratio	95% Confidence Interval	
		Lower	Upper					Lower	Upper
Na_Onset	-0.113	-0.194	-0.0329	0.0409	-2.77	0.006	0.893	0.824	0.967

Tab. 16 – PDPs for creatinine

Sodium

Adjusting for creatinine, a significant negative relation between serum Na and risk of death was confirmed by regression model (already reported above, Tab 16). PDPs show a clear rising in risk when Na levels are between 120 and 130 mg/dl (Fig.18).

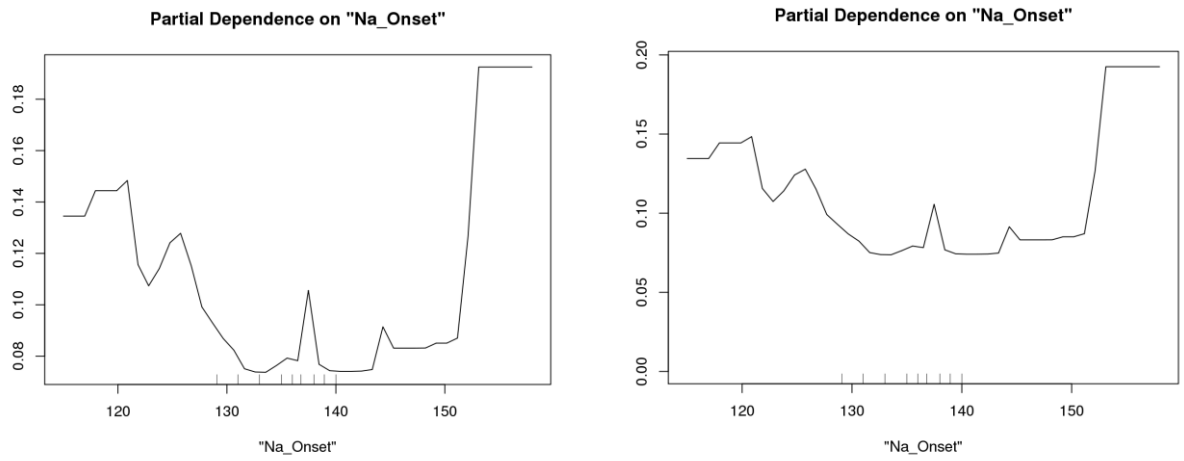


Fig. 18 – PDPs for sodium

This effect does not depend on a relationship between creatinine and Na (possible due to AKI), as confirmed by mediation analysis (Tab 17).

Na-> Creatinine->Death

			95% Confidence Interval				
Effect	**Estimate**	**SE**	**Lower**	**Upper**	**Z**	**p**	**% Mediation**
Indirect	-3.48e−4	8.15e-4	-0.00200	0.00164	-0.427	0.669	3.69
Direct	-0.00910	0.00508	-0.01946	-3.44e−4	-1.791	0.073	96.31
Total	-0.00944	0.00526	-0.02023	1.82e-4	-1.796	0.072	100.00

Tab. 17 – Mediation analysis for the path Na-> Creatinine->Death

We also assessed if the effect of hyponatriemia on mortality could be explained by the assosiation between death risk and the underlying hyperinflammation (CRP levels) or by hemodinamic changes due to multiorgan failure (MOF). The mediation analyses showed no indirect effect in both cases (Tab. 18 and 19).

CRP->Na-> Death model

			95% Confidence Interval				
Effect	**Estimate**	**SE**	**Lower**	**Upper**	**Z**	**p**	**% Mediation**
Indirect	-7.12e−4	8.51e-4	-0.00261	9.71e-4	-0.837	0.403	7.54
Direct	-0.00873	0.00489	-0.01841	2.86e-4	-1.785	0.074	92.46
Total	-0.00944	0.00511	-0.01979	2.23e-4	-1.849	0.064	100.00

Tab. 18 – Mediation analysis for the path CRP -> Na-> Death

MOF->Na->Death

			95% Confidence Interval				
Effect	**Estimate**	**SE**	**Lower**	**Upper**	**Z**	**p**	**% Mediation**
Indirect	0.00928	0.0129	-0.00586	0.0414	0.719	0.472	5.40
Direct	0.16252	0.0768	0.00597	0.3188	2.115	0.034	94.60
Total	0.17180	0.0742	0.02762	0.3299	2.314	0.021	100.00

Tab. 19– Mediation analysis for the path MOF -> Na-> Death

LDH

The adjustment set included creatinine and sodium. As shown in PDPs (Fig.19), a slight significant positive correlation (Spearman's r 0.344, p-value 0.013) emerged between LDH levels and mortality, albeit LDH was not an indipendent predictor in the regression model (Tab. 20)

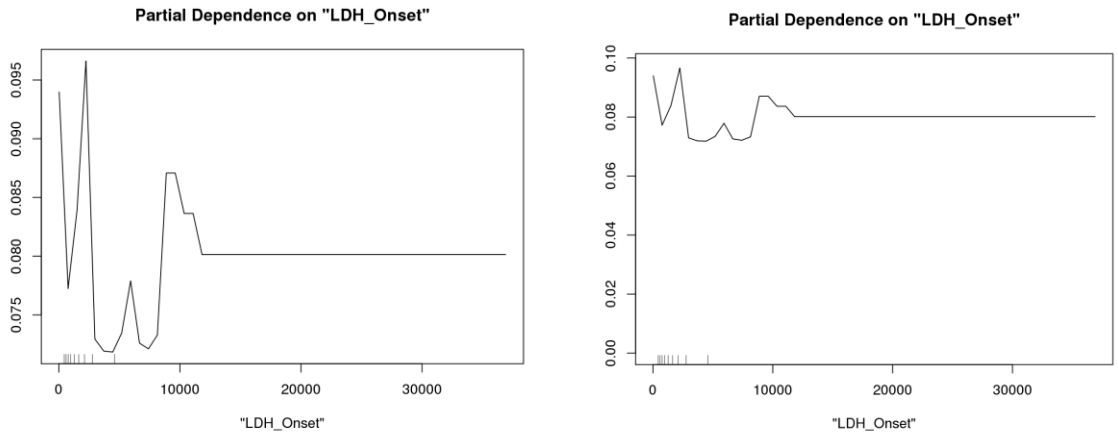


Fig 19– PDPs for LDH

		95% Confidence Interval						95% Confidence Interval	
Predictor	**Estimate**	**Lower**	**Upper**	**SE**	**Z**	**p**	**Odds ratio**	**Lower**	**Upper**
Intercept	12.672	1.913	23.4308	5.4892	2.309	0.021	318727.812	6.776	1.50e+10
Crea_Onset	0.592	0.144	1.0394	0.2283	2.593	0.010	1.807	1.155	2.828

		95% Confidence Interval						95% Confidence Interval	
Predictor	**Estimate**	**Lower**	**Upper**	**SE**	**Z**	**p**	**Odds ratio**	**Lower**	**Upper**
Na_Onset	-0.115	-0.195	-0.0346	0.0410	-2.804	0.005	0.891	0.823	0.966
LDH_Onset	-6.93e−5	-2.22e−4	8.35e-5	7.79e-5	-0.889	0.374	1.000	1.000	1.000

Tab. 20– regression model for LDH

Mediation modelling for the association between serum creatinine values (AKI) and risk of Death

To investigate the possible mechanism underlying the causal pathway linking incremented creatinine values (AKI) to mortality, we compared different explanatory models through causal mediation analysis. The conceptual models, based on existing knowledge of potential mechanisms of kindey injury in HLH diseases[72], are illustrated in Fig.20.

[72] Karras, Alexandre. "What nephrologists need to know about hemophagocytic syndrome." Nature Reviews Nephrology 5.6 (2009): 329.

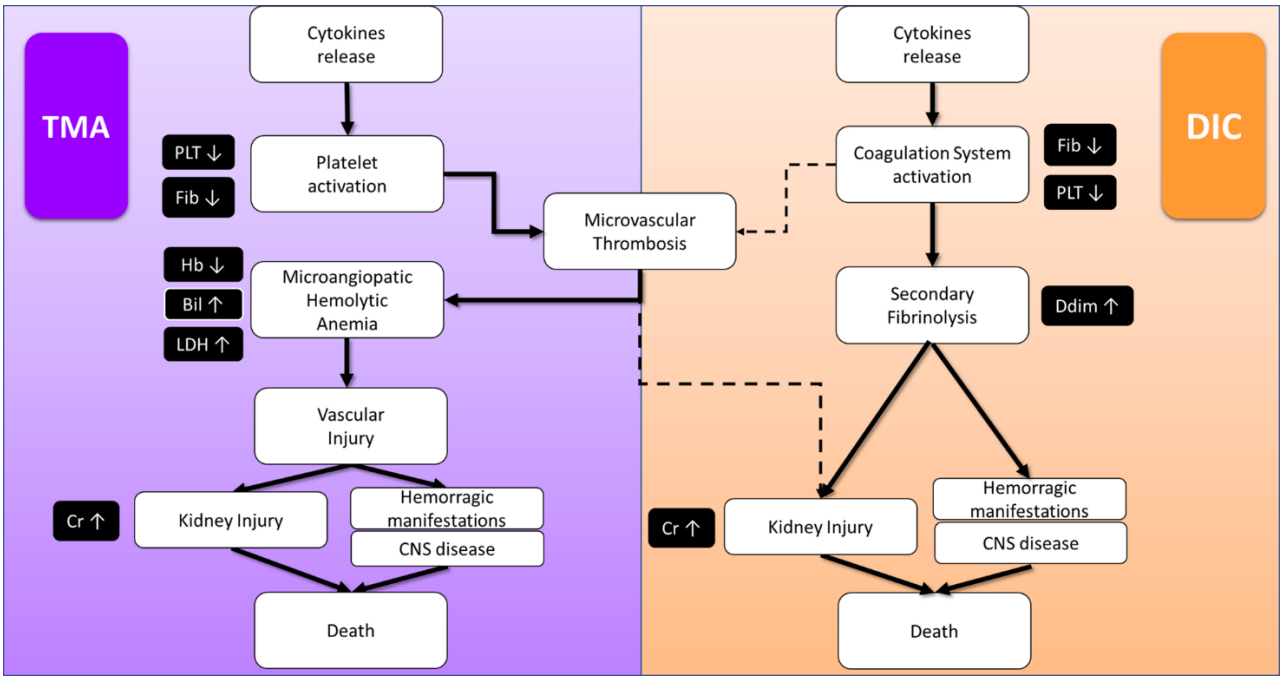


Fig.20 – Models of two potential pathofisiological mechanisms driving AKI in MAS, DIC-like coagulopahty (disseminated intravascular coagulation) and TMA (thrombotic microangiopahy). Measurable parameters useful to empirically test the two models are depicted in black boxes. Modified from Wada, Hideo, et al. "Differences and similarities between disseminated intravascular coagulation and thrombotic microangiopathy." Thrombosis journal 16.1 (2018): 14.

TMA model

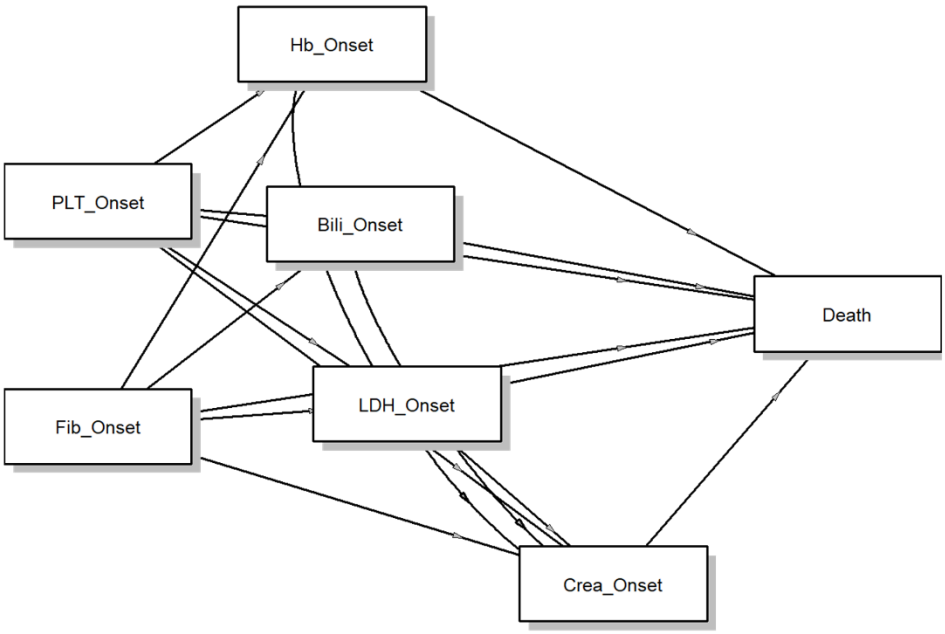


Fig.21 –The mediation model formalizing the hypothesis of a TMA-induced kidney injury.

Fig. 21 shows a DAG representing the possible dependencies between laboratory variables based on known mechanisms uderlying thrombotic microangiopahy (TMA), a recognized complication of MAS[73]. To assess if such a mechanism is able to explain the emerged assosiation between AKI and death in our cohort, a generalized mediation model was fitted. Results, reported in Tab. 21, did not support a pathway involving PLT -> Hb -> Bil. And LDH -> creatinine -> death, suggesting that that microangiopathic hemolytic disease is not the main pathogenetic process leading to fatal outcomes in our sample.

[73] Minoia F, Tibaldi J, Muratore V, et alFRI0565 A MULTINATIONAL STUDY OF THROMBOTIC MICROANGIOPATHY IN MACROPHAGE ACTIVATION SYNDROME: A DREADFUL CONDITION WHICH IS LIKELY UNDER-RECOGNIZEDAnnals of the Rheumatic Diseases 2019;78:978.

Tab. 21 -TMA mediation model

Type	Effect	Estimate	SE	95% C.I. (a)		β	z	p
				Lower	Upper			
Indirect	PLT_Onset ⇒ Hb_Onset ⇒ Death	2.56e-7	4.72e-6	-9.01e−6	9.48e-6	1.88e-4	0.0543	0.957
	PLT_Onset ⇒ Bili_Onset ⇒ Death	2.14e-6	9.15e-6	-1.62e−5	1.97e-5	0.00157	0.2336	0.815
	PLT_Onset ⇒ LDH_Onset ⇒ Death	6.64e-6	7.11e-6	-7.38e−6	2.05e-5	0.00486	0.9339	0.350
	PLT_Onset ⇒ Crea_Onset ⇒ Death	-2.15e−5	1.61e-5	-5.34e−5	9.68e-6	-0.01576	-1.3375	0.181
	Fib_Onset ⇒ Hb_Onset ⇒ Death	-3.78e−6	1.06e-5	-2.35e−5	1.80e-5	-0.00277	-0.3571	0.721

Tab. 21 -TMA mediation model

Type	Effect	Estimate	SE	95% C.I. (a) Lower	95% C.I. (a) Upper	β	z	p
	Fib_Onset ⇒ Bili_Onset ⇒ Death	-1.03e−5	1.67e-5	-4.28e−5	2.28e-5	-0.00753	-0.6135	0.540
	Fib_Onset ⇒ LDH_Onset ⇒ Death	2.78e-5	1.37e-5	-3.90e−7	5.34e-5	0.02040	2.0267	0.043
	Fib_Onset ⇒ Crea_Onset ⇒ Death	1.11e-5	1.89e-5	-2.58e−5	4.82e-5	0.00812	0.5873	0.557
	PLT_Onset ⇒ Hb_Onset ⇒ Crea_Onset ⇒ Death	9.45e-8	9.07e-7	-1.67e−6	1.88e-6	6.92e-5	0.1041	0.917

Tab. 21 -TMA mediation model

Type	Effect	Estimate	SE	95% C.I. (a) Lower	95% C.I. (a) Upper	β	z	p
	PLT_Onset ⇒ Bili_Onset ⇒ Crea_Onset ⇒ Death	-6.30e−8	1.13e-6	-2.28e−6	2.14e-6	-4.61e−5	-0.055	0.956
	PLT_Onset ⇒ LDH_Onset ⇒ Crea_Onset ⇒ Death	-3.20e−6	3.92e-6	-1.08e−5	4.56e-6	-0.0023	-0.815	0.415
	Fib_Onset ⇒ Hb_Onset ⇒ Crea_Onset ⇒ Death	-1.39e−6	2.29e-6	-5.85e−6	3.13e-6	-0.0010	-0.608	0.543

Tab. 21 -TMA mediation model

Type	Effect	Estimate	SE	95% C.I. (a) Lower	95% C.I. (a) Upper	β	z	p
	Fib_Onset ⇒ Bili_Onset ⇒ Crea_Onset ⇒ Death	3.03e-7	2.04e-6	-3.64e−6	4.37e-6	2.22e-4	0.1480	0.882
	Fib_Onset ⇒ LDH_Onset ⇒ Crea_Onset ⇒ Death	-1.34e−5	8.66e-6	-2.96e−5	4.34e-6	-0.00982	-1.5477	0.122
Component	PLT_Onset ⇒ Hb_Onset	-8.50e−5	5.45e-4	-0.00117	9.70e-4	-0.00802	-0.1560	0.876
	Hb_Onset ⇒ Death	-0.00301	0.00758	-0.01706	0.01267	-0.02339	-0.3974	0.691
	PLT_Onset ⇒ Bili_Onset	3.59e-4	9.47e-4	-0.00156	0.00215	0.03240	0.3795	0.704

Tab. 21 -TMA mediation model

Type	Effect	Estimate	SE	95% C.I. (a)		β	z	p
				Lower	Upper			
	Bili_Onset ⇒ Death	0.00595	0.00846	-0.01068	0.02249	0.04831	0.7033	0.482
	PLT_Onset ⇒ LDH_Onset	-0.86520	0.74607	-2.34842	0.57613	-0.05035	-1.1597	0.246
	LDH_Onset ⇒ Death	-7.67e−6	3.46e-6	-1.42e−5	-5.99e−7	-0.09653	-2.2169	0.027
	PLT_Onset ⇒ Crea_Onset	-2.95e−4	1.38e-4	-5.70e−4	-2.97e−5	-0.09166	-2.1400	0.032
	Crea_Onset ⇒ Death	0.07301	0.03805	-0.00165	0.14751	0.17193	1.9187	0.055

Tab. 21 -TMA mediation model

Type	Effect	Estimate	SE	95% C.I. (a) Lower	95% C.I. (a) Upper	β	z	p
	Fib_Onset ⇒ Hb_Onset	0.00125	5.72e-4	1.38e-4	0.00238	0.11841	2.1926	0.028
	Fib_Onset ⇒ Bili_Onset	-0.00173	8.95e-4	-0.00345	5.84e-5	-0.15587	-1.9284	0.054
	Fib_Onset ⇒ LDH_Onset	-3.62787	0.93886	-5.43624	-1.75598	-0.21133	-3.8641	< .001
	Fib_Onset ⇒ Crea_Onset	1.52e-4	2.34e-4	-3.05e−4	6.12e-4	0.04725	0.6488	0.516
	Hb_Onset ⇒ Crea_Onset	-0.01522	0.02162	-0.05718	0.02758	-0.05017	-0.7039	0.482

Tab. 21 -TMA mediation model

Type	Effect	Estimate	SE	95% C.I. (a)		β	z	p
				Lower	Upper			
	Bili_Onset ⇒ Crea_Onset	-0.00240	0.01293	-0.02827	0.02242	-0.00828	-0.1857	0.853
	LDH_Onset ⇒ Crea_Onset	5.06e-5	2.16e-5	5.86e-6	9.06e-5	0.27033	2.3416	0.019
Direct	PLT_Onset ⇒ Death	-5.46e−5	7.05e-5	-1.92e−4	8.42e-5	-0.03997	-0.7746	0.439
	Fib_Onset ⇒ Death	-1.43e−4	8.85e-5	-3.19e−4	2.77e-5	-0.10446	-1.6106	0.107
Total	PLT_Onset ⇒ Death	-7.02e−5	7.88e-5	-2.25e−4	8.41e-5	-0.05147	-0.8919	0.372

Tab. 21 -TMA mediation model

Type	Effect	Estimate	SE	95% C.I. (a) Lower	95% C.I. (a) Upper	β	z	p
	Fib_Onset ⇒ Death	-1.32e−4	7.87e-5	-2.86e−4	2.21e-5	-0.0969	-1.679	0.093

DIC-like coagulophaty model

The laboratory relationships assumed in the model supposing a causal role for disseminated intravascular-like coagulophaty are decipted in Fig. 22.

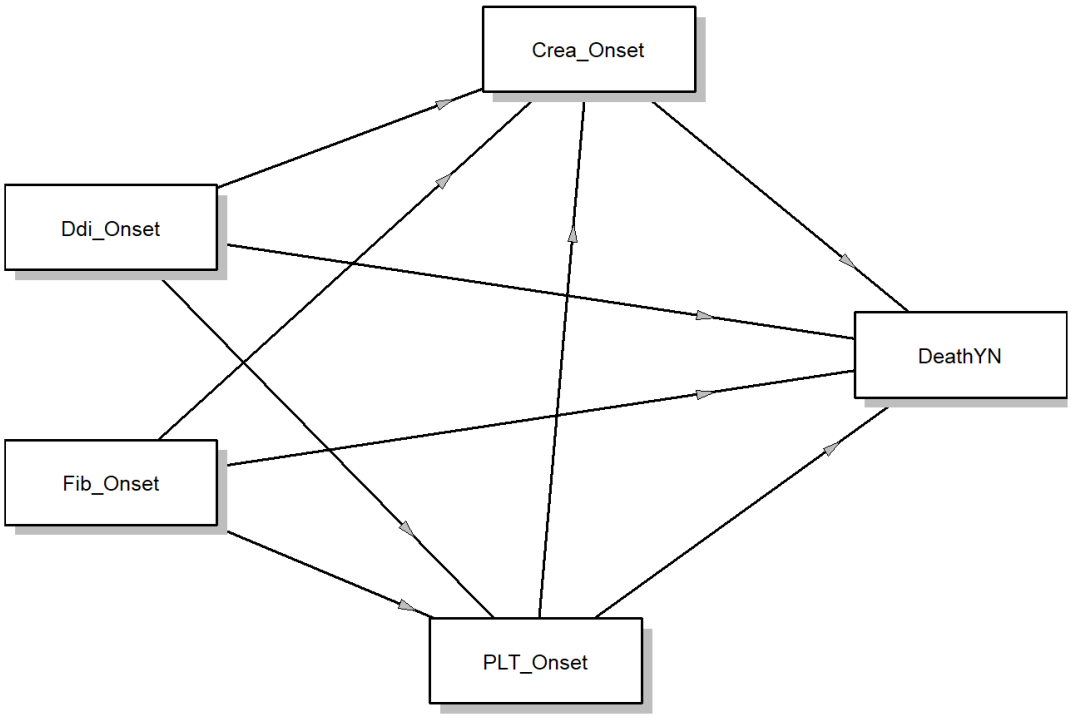


Fig 22 – DAG for a DIC-mediated kidney injury in SJIA-MAS. Covariance between indipendent variables are not shown

As reported in Tab.22, a significant indirect effect links d-dimer values (compsuntive coagulopathy) and death risk, through kidney dysfunction (creatinine). As d-dimer has not direct effect on mortality, its effect appears totally mediated by kidney injury. While fibrinogen is associated with PLT levels; however, indirect effects involving these two parameters are not significant.

Tab. 22 -DIC-like coagulophaty model

Type	Effect	Estimate	SE	95% C.I. (a) Lower	95% C.I. (a) Upper	β	z	p
Indirect	Ddi_Onset ⇒ Crea_Onset ⇒ DeathYN	0.00978	0.00423	0.00149	0.01807	0.03520	2.313	0.021
	Ddi_Onset ⇒ PLT_Onset ⇒ DeathYN	6.70e-4	0.00122	-0.00172	0.00306	0.00241	0.549	0.583
	Fib_Onset ⇒ Crea_Onset ⇒ DeathYN	-5.76e−4	0.00223	-0.00494	0.00379	-0.00210	-0.259	0.796
	Fib_Onset ⇒ PLT_Onset ⇒ DeathYN	-0.00360	0.00607	-0.01550	0.00831	-0.01314	-0.592	0.554
	Ddi_Onset ⇒ PLT_Onset ⇒ Crea_Onset ⇒ DeathYN	2.42e-4	2.49e-4	-2.46e−4	7.30e-4	8.69e-4	0.970	0.332

Tab. 22 -DIC-like coagulophaty model

Type	Effect	Estimate	SE	95% C.I. (a) Lower	95% C.I. (a) Upper	β	z	p
Component	Fib_Onset ⇒ PLT_Onset ⇒ Crea_Onset ⇒ DeathYN	-0.00130	0.00100	-0.00326	6.65e-4	-0.00474	-1.295	0.195
	Ddi_Onset ⇒ Crea_Onset	0.25216	0.05362	0.14706	0.35726	0.24342	4.702	< .001
	Crea_Onset ⇒ DeathYN	0.03879	0.01461	0.01016	0.06741	0.14460	2.656	0.008
	Ddi_Onset ⇒ PLT_Onset	-0.07281	0.05056	-0.17190	0.02628	-0.07086	-1.440	0.150
	PLT_Onset ⇒ DeathYN	-0.00920	0.01550	-0.03957	0.02117	-0.03403	-0.594	0.553

Tab. 22 -DIC-like coagulophaty model

Type	Effect	Estimate	SE	95% C.I. (a) Lower	95% C.I. (a) Upper	β	z	p
	Fib_Onset ⇒ Crea_Onset	-0.01486	0.05714	-0.12684	0.09713	-0.01456	-0.260	0.795
	Fib_Onset ⇒ PLT_Onset	0.39091	0.04981	0.29328	0.48854	0.38615	7.848	< .001
	PLT_Onset ⇒ Crea_Onset	-0.08554	0.05661	-0.19649	0.02540	-0.08485	-1.511	0.131
Direct	Ddi_Onset ⇒ DeathYN	0.00746	0.01509	-0.02211	0.03703	0.02685	0.494	0.621
	Fib_Onset ⇒ DeathYN	-0.02596	0.01559	-0.05652	0.00460	-0.09482	-1.665	0.096

Tab. 22 -DIC-like coagulophaty model

Type	Effect	Estimate	SE	95% C.I. (a) Lower	95% C.I. (a) Upper	β	z	p
Total	Ddi_Onset ⇒ DeathYN	0.01815	0.01477	-0.01079	0.04710	0.06533	1.229	0.219
	Fib_Onset ⇒ DeathYN	-0.03143	0.01455	-0.05995	0.00291	-0.11480	-2.160	0.031

Demographic features

Age at sJIA onset

With an adjustment set including AST, age at MAS onset, creatinine, ddimer, fibrinogen, and LDH, we found not significant association with mortality.

Model Coefficients

Predictor	Estimate	95% Confidence Interval		SE	Z	p	Odds ratio	95% Confidence Interval	
		Lower	Upper					Lower	Upper
Intercept	-2.22588	-3.33305	-1.11870	0.56490	-3.940	< .001	0.108	0.0357	0.327
Fib_Onset	-0.00281	-0.00545	-1.75e−4	0.00135	-2.090	0.037	0.997	0.9946	1.000
LDH_Onset	-1.58e−4	-3.47e−4	3.08e-5	9.64e-5	-1.641	0.101	1.000	0.9997	1.000
Age_at_MAS_onset	0.04926	-0.05473	0.15323	0.05305	0.928	0.353	1.050	0.9467	1.166
AST_Onset	2.31e-4	-7.00e−5	5.32e-4	1.53e-4	1.504	0.133	1.000	0.9999	1.001

Model Coefficients

Predictor	Estimate	95% Confidence Interval		SE	Z	p	Odds ratio	95% Confidence Interval	
		Lower	Upper					Lower	Upper
Ddi_Onset	2.62e-5	-3.06e−5	8.30e-5	2.90e-5	0.904	0.366	1.000	1.0000	1.000
Crea_Onset	0.63471	0.17234	1.0971	0.23591	2.691	0.007	1.886	1.1881	2.995
Age_at_onset	-0.05048	-0.17370	0.0727	-0.06287	-0.803	0.422	0.951	0.8405	1.075

Tab. 23 – regression for Age at SJIA onset

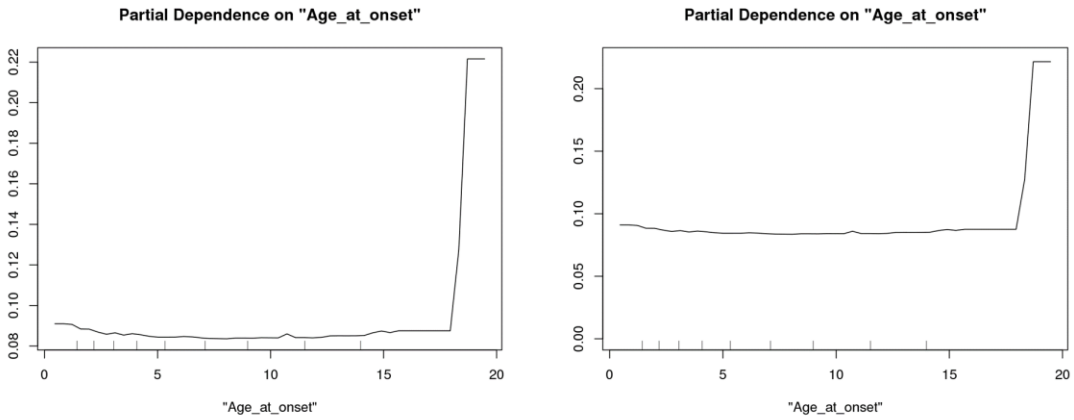


Fig. 23 – PDPs for Age at SJIA onset

Age at MAS onset

With an adjustment set including creatinine, ddimer, fibrinogen, Hb, LDH and sodium values, no significant relationship with mortality emerged from linear and MARS models, despite an apparent positive correlation, illustrated in PDPs (Spearman's rank correlation of 0.949, p-value 0.000).

Tab. 24 – regression model for age at MAS onset

Predictor	Estimate	95% Confidence Interval		SE	Z	p	Odds ratio	95% Confidence Interval	
		Lower	Upper					Lower	Upper
Intercept	10.62974	-0.37126	21.6307	5.61286	1.8938	0.058	41346.208	0.690	2.48e+9
Fib_Onset	-0.00272	-0.00531	-1.24e−4	0.00132	-2.0533	0.040	0.997	0.995	1.000
LDH_Onset	-1.25e−4	-3.02e−4	5.29e-5	9.06e-5	-1.3766	0.169	1.000	1.000	1.000
Age_at_MAS_onset	0.06074	-0.04935	0.1708	0.05617	1.0814	0.280	1.063	0.952	1.186
Ddi_Onset	2.33e-5	-3.27e−5	7.93e-5	2.86e-5	0.8146	0.415	1.000	1.000	1.000

Tab. 24 – regression model for age at MAS onset

Predictor	Estimate	95% Confidence Interval		SE	Z	p	Odds ratio	95% Confidence Interval	
		Lower	Upper					Lower	Upper
Crea_Onset	0.57347	0.09061	1.0563	0.24636	2.3278	0.020	1.774	1.095	2.876
Age_at_onset	-0.05215	-0.17452	0.0702	0.06244	-0.8352	0.404	0.949	0.840	1.073
Hb_Onset	-0.00940	-0.20942	0.1906	0.10205	-0.0921	0.927	0.991	0.811	1.210
Na_Onset	-0.09512	-0.17901	-0.0112	0.04280	-2.2225	0.026	0.909	0.836	0.989

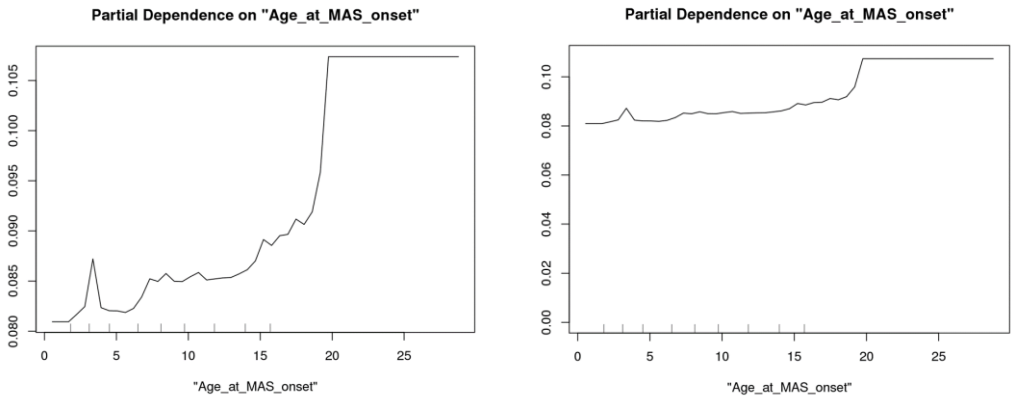


Fig. 24 – PDPs for Age at MAS onset

Added-variable plots for death and selected parameters are shown below (Fig. 25)

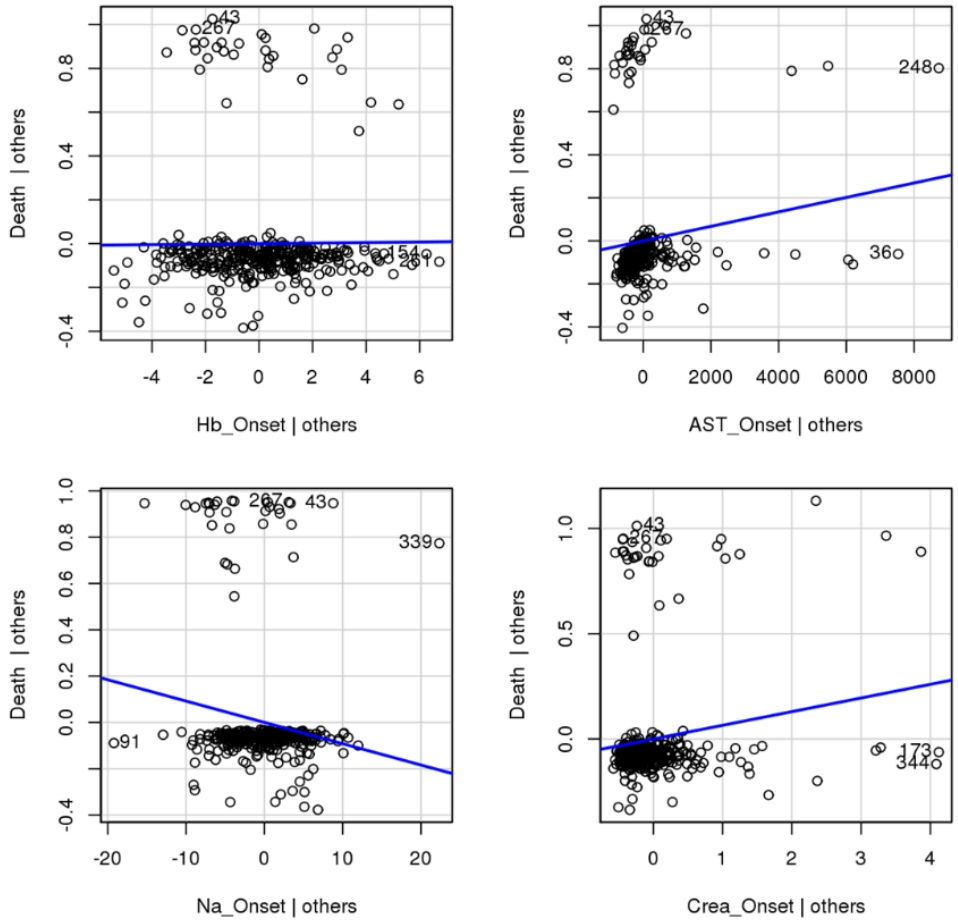


Fig.25 – Added-variable plots summarizing the relatioship between death and selected parameters

3.1.3 CNS disease

The DAG and BN for CNS disease are reported below (Fig. 26)

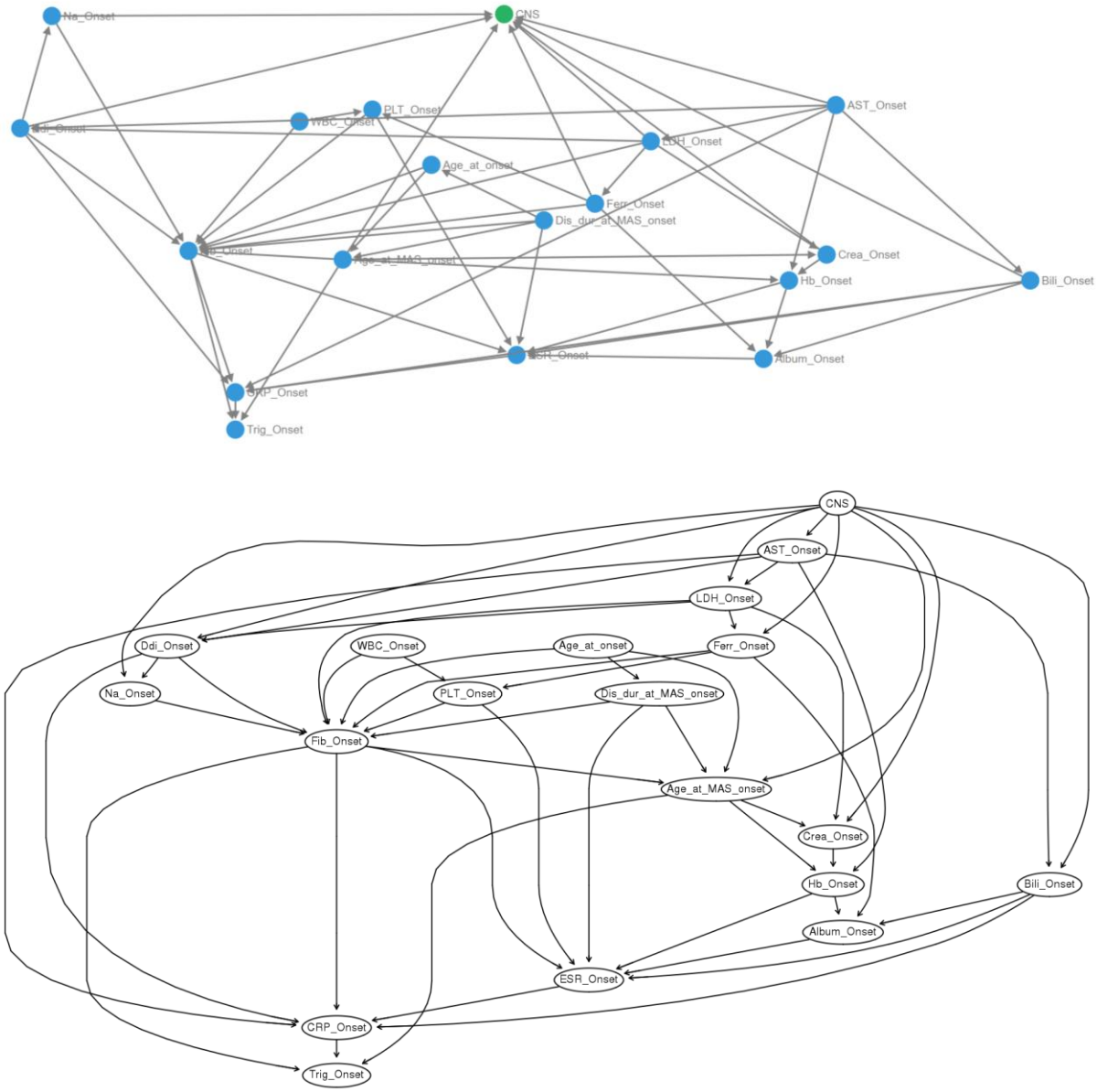


Fig.26 – A priori and data-driven DAGs for CNS disease

Age at MAS onset

When adjusting for LDH and creatinine, a negative correlation between age at MAS onset and CNS was observed, based on Spearman's rank correlation of -0.685, p-value 0.000 (PDPs shown in Fig.27). However, logistic regression did not show a significant association (Tab. 25)

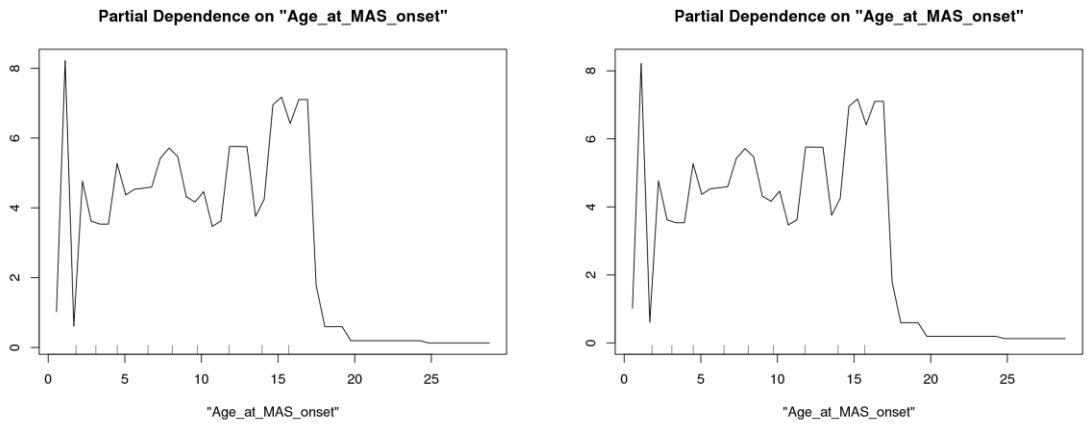


Fig.27 – PDPs for age at MAS onset

Predictor	Estimate	95% Confidence Interval		SE	Z	p	Odds ratio	95% Confidence Interval	
		Lower	Upper					Lower	Upper
Intercept	-1.22283	-1.73663	-0.70903	0.2620	-4.668	< .001	0.294	0.176	0.492

Predictor	Estimate	95% Confidence Interval		SE	Z	p	Odds ratio	95% Confidence Interval	
		Lower	Upper					Lower	Upper
Age_at_MAS_onset	-0.0163	-0.0617	0.0291	0.0231	-0.704	0.482	0.984	0.940	1.029
LDH_Onset	8.97e-5	3.83e-6	1.76e-4	4.38e-5	2.047	0.041	1.000	1.000	1.000
Crea_Onset	0.8032	0.3069	1.2994	0.2532	3.172	0.002	2.233	1.359	3.667

Note. Estimates represent the log odds of "CNS = 1" vs. "CNS = 0"

Tab. 25 – regression model for age at MAS onset

Age at sJIA onset

Adjusting for age at MAS onset, there was a negative correlation (Spearman's r -0.447, p-value 0.001) between age at sJIA onset and CNS dysfunction (PDPs in Fig. 28). No significant relationship was detected by regression modelling (Tab. 26).

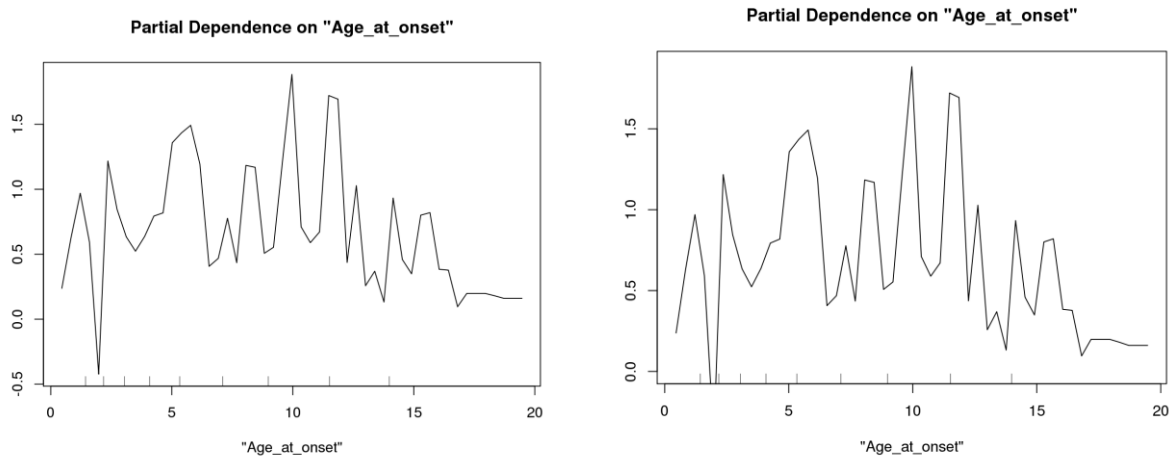


Fig.28 – PDPs for age at sJIA onset

Predictor	Estimate	95% Confidence Interval		SE	Z	p	Odds ratio	95% Confidence Interval	
		Lower	Upper					Lower	Upper
Intercept	-0.63115	-1.0542	-0.2081	0.2158	-2.9242	0.003	0.532	0.348	0.812
Age_at_MAS_onset	0.00323	-0.0645	0.0709	0.0345	0.0934	0.926	1.003	0.938	1.073

Predictor	Estimate	95% Confidence Interval		SE	Z	p	Odds ratio	95% Confidence Interval	
		Lower	Upper					Lower	Upper
Age_at_onset	-0.00261	-0.07879	0.0737	0.0389	-0.0669	0.947	0.997	0.924	1.077

Tab. 25 – regression model for age at sJIA onset

Disease duration

Despite the PDPs show a significantly negative (based on Spearman's rank correlation of -0.923 and p-value 0.000) after adjustment for age at MAS onset, no evidence of association was found with the regression model (Tab 26)

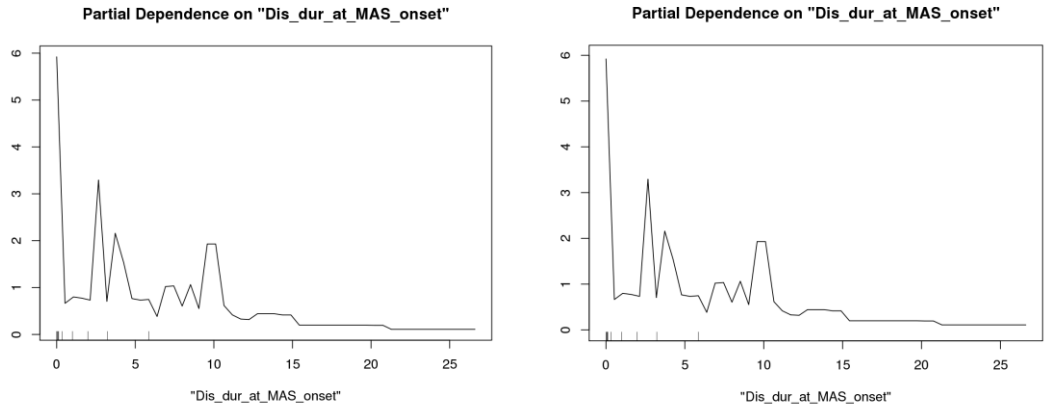


Fig.29 – PDPs for age disease duration

Predictor	Estimate	95% Confidence Interval		SE	Z	p	Odds ratio	95% Confidence Interval	
		Lower	Upper					Lower	Upper
Intercept	-0.62337	-1.0451	-0.2016	0.2152	-2.897	0.004	0.536	0.352	0.817
Age_at_MAS_onset	-0.00358	-0.0516	0.0444	0.0245	-0.146	0.884	0.996	0.950	1.045

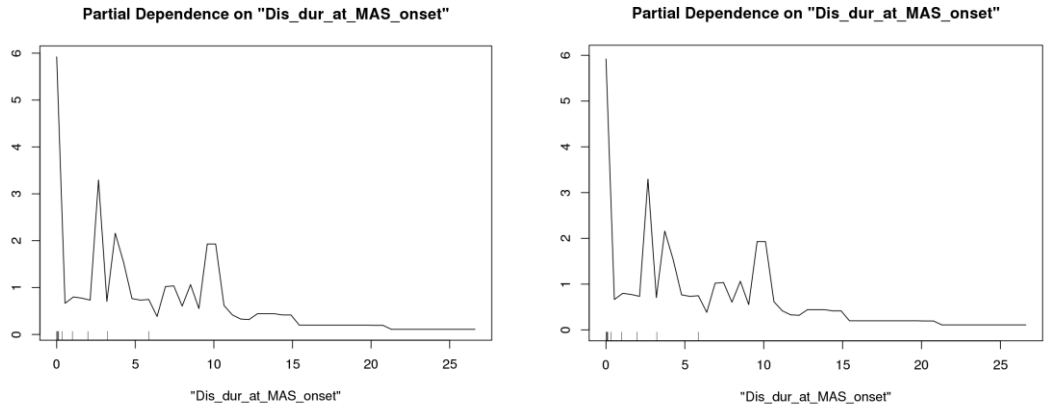


Fig.29 – PDPs for age disease duration

Predictor	Estimate	95% Confidence Interval		SE	Z	p	Odds ratio	95% Confidence Interval	
		Lower	Upper					Lower	Upper
Dis_dur_at_MAS_onset	0.01652	-0.0576	0.0906	0.0378	0.437	0.662	1.017	0.944	1.095

Tab. 26 – regression model for disease duration

Fibrinogen

No adjustment set was required to assess the direct effect of fibrinogen levels (coagulopathy) on CNS disease. A significant association was confirmed by logistic regression (Tab. 27), with higher values being protective for CNS involvement development. We further assessed if this relationship depends on the occurrence of clinical hemorrhagic manifestations, through a mediation analysis. As shown in Table 27 and Fig.30, we found that the impact of fibrinogen is only partially (25% of the effect) explained by hemorrhagic complications.

Model Coefficients

		95% Confidence Interval					95% Confidence Interval	
Predictor	**Estimate**	**Lower**	**Upper**	**SE**	**Z**	**p**	**Odds ratio**	**Lower**	**Upper**
Intercept	0.06621	-0.35210	0.4845	0.213	0.310	0.756	1.068	0.703	1.6223
Fib_Onset	-0.00231	-0.00357	-0.0010	6.40e-4	-3.611	< .001	0.998	0.996	0.999

Model Coefficients

Predictor	Estimate	95% Confidence Interval		SE	Z	p	Odds ratio	95% Confidence Interval	
		Lower	Upper					Lower	Upper

Mediation Estimates

Effect	Estimate	SE	95% Confidence Interval		Z	p	% Mediation
			Lower	Upper			
Indirect	-1.14e−4	4.27e-5	-2.13e−4	-4.52e−5	-2.67	0.008	25.3
Direct	-3.37e−4	1.37e-4	-6.01e−4	-6.10e−5	-2.46	0.014	74.7
Total	-4.51e−4	1.38e-4	-7.25e−4	-1.73e−4	-3.27	0.001	100.0

Path Estimates

					95% Confidence Interval			
			Estimate	SE	Lower	Upper	Z	p
Fib_Onset	→	Hemorrhagic	-3.73e−4	1.01e-4	-5.79e−4	-1.81e−4	-3.68	< .001
Hemorrhagic	→	CNS	0.305	0.0660	0.170	0.436	4.63	< .001
Fib_Onset	→	CNS	-3.37e−4	1.37e-4	-6.01e−4	-6.10e−5	-2.46	0.014

Tab. 27 – regression model for fibrinogen and mediation model for the path fibrinogen-> hemorragic manifestations-> CNS disease

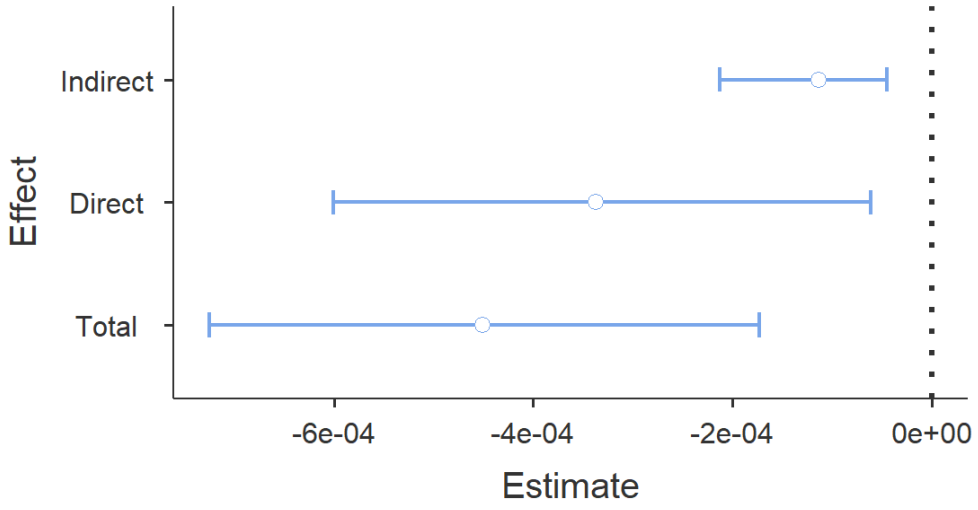


Fig.30 – Direct and indirect effects in the mediation model for the path fibrinogen-> hemorragic manifestations-> CNS disease

Ferritin

No adjustment was need according to the DAG. Ferritin levels were significantly associated with CNS disease occurrence (OR 1.69, 95%Cl 1.33 – 2.1). A significant interaction effect emerged between ferritin and age at MAS onset, such as for specific ferritin values the risk of CNS involvement is higher at younger age, as shown in Tab. 28 and Fig. 31.

		95% Confidence Interval						95% Confidence Interval	
Predictor	**Estimate**	**Lower**	**Upper**	**SE**	**Z**	**p**	**Odds ratio**	**Lower**	**Upper**
Intercept	-0.636	-0.864	-0.409	0.116	-5.48	< .001	0.529	0.422	0.664
Ferr_Onset	0.530	0.291	0.769	0.122	4.34	< .001	1.699	1.338	2.158

Note. Estimates represent the log odds of "CNS = 1" vs. "CNS = 0"

			95% Confidence Interval			
	Estimate	**SE**	**Lower**	**Upper**	**Z**	**p**
Ferr_Onset	0.13202	0.02496	0.0831	0.18095	5.289	< .001
Age_at_MAS_onset	-0.00192	0.00472	-0.0112	0.00733	-0.406	0.684
Ferr_Onset ✻ Age_at_MAS_onset	-0.01033	0.00512	-0.0204	2.96e−4	-2.018	0.044

			95% Confidence Interval			
	Estimate	SE	Lower	Upper	Z	p

			95% Confidence Interval			
	Estimate	SE	Lower	Upper	Z	p
Average	0.1328	0.0252	0.0834	0.182	5.27	< .001
Low (-1SD)	0.1865	0.0408	0.1066	0.266	4.58	< .001
High (+1SD)	0.0791	0.0322	0.0161	0.142	2.46	0.014

Tab. 28 – Moderation model forthe effect of ferritin levels on CNS disease. The table shows the effect of the predictor (ferritin) on the dependent variable (CNS) at different levels of the moderator (age at MAS onset)

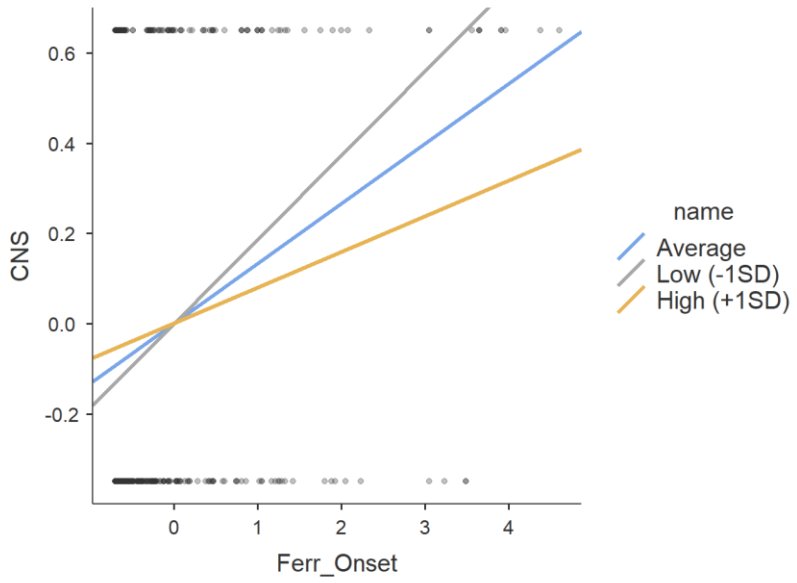


Fig.31 – The plot illustrates the relatioship netween ferritin values and risk of CNS disease according to different values of age at MAS onset (colored slopes)

Creatinine

In absence of a required adjustment set, the estimated directed effect of creatinine levels (AKI) on CNS disease occurrence was significant (Tab.29). Mediation analysis showed that this association is partly (around 20%) due to the association of AKI with heart failure (Tab 31 and Fig. 32).

Model Coefficients

		95% Confidence Interval						95% Confidence Interval	
Predictor	**Estimate**	**Lower**	**Upper**	**SE**	**Z**	**p**	**Odds ratio**	**Lower**	**Upper**
Intercept	-1.223	-1.610	-0.837	0.197	-6.20	< .001	0.294	0.200	0.433
Crea_Onset	0.889	0.410	1.369	0.245	3.63	< .001	2.434	1.506	3.932

Tab. 29 – Regression model of creatinine for CNS disease

Mediation Estimates

			95% Confidence Interval				
Effect	**Estimate**	**SE**	**Lower**	**Upper**	**Z**	**p**	**% Mediation**
Indirect	0.0319	0.0153	0.00139	0.0645	2.08	0.037	17.6
Direct	0.1492	0.0435	0.07023	0.2408	3.43	< .001	82.4
Total	0.1811	0.0386	0.10498	0.2606	4.69	< .001	100.0

Path Estimates

					95% Confidence Interval			
			Estimate	SE	Lower	Upper	Z	p
Crea_Onset	→	heartfail	0.0753	0.0341	0.0101	0.141	2.21	0.027
heartfail	→	CNS	0.4234	0.1287	0.1141	0.627	3.29	0.001
Crea_Onset	→	CNS	0.1492	0.0435	0.0702	0.241	3.43	< .001

Tab. 31 – Mediation model for the path cratinine-> heart failure-> CNS

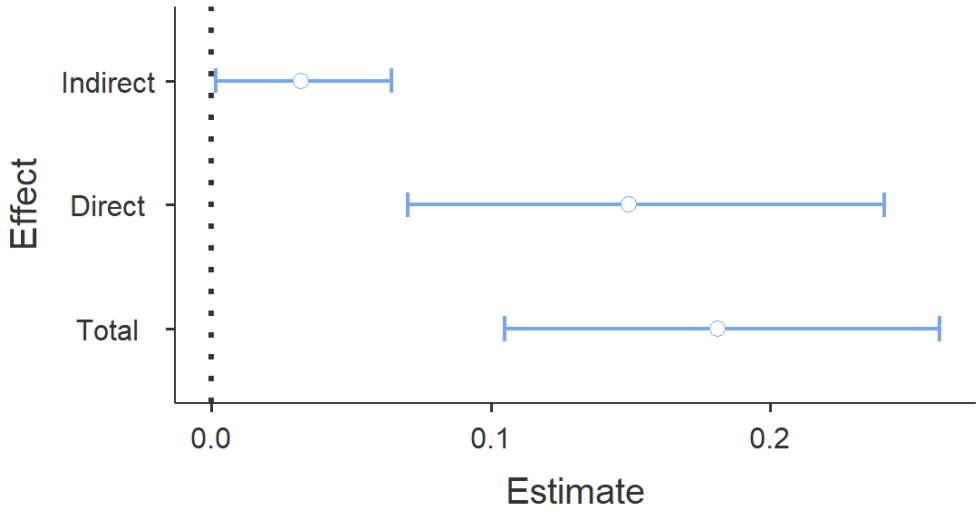


Fig. 32 –Direct and indirect effects in r the path cratinine-> heart failure-> CNS

Added-variable plots highlight the relationship between the risk of CNS disease and the examined variables (Fig. 33)

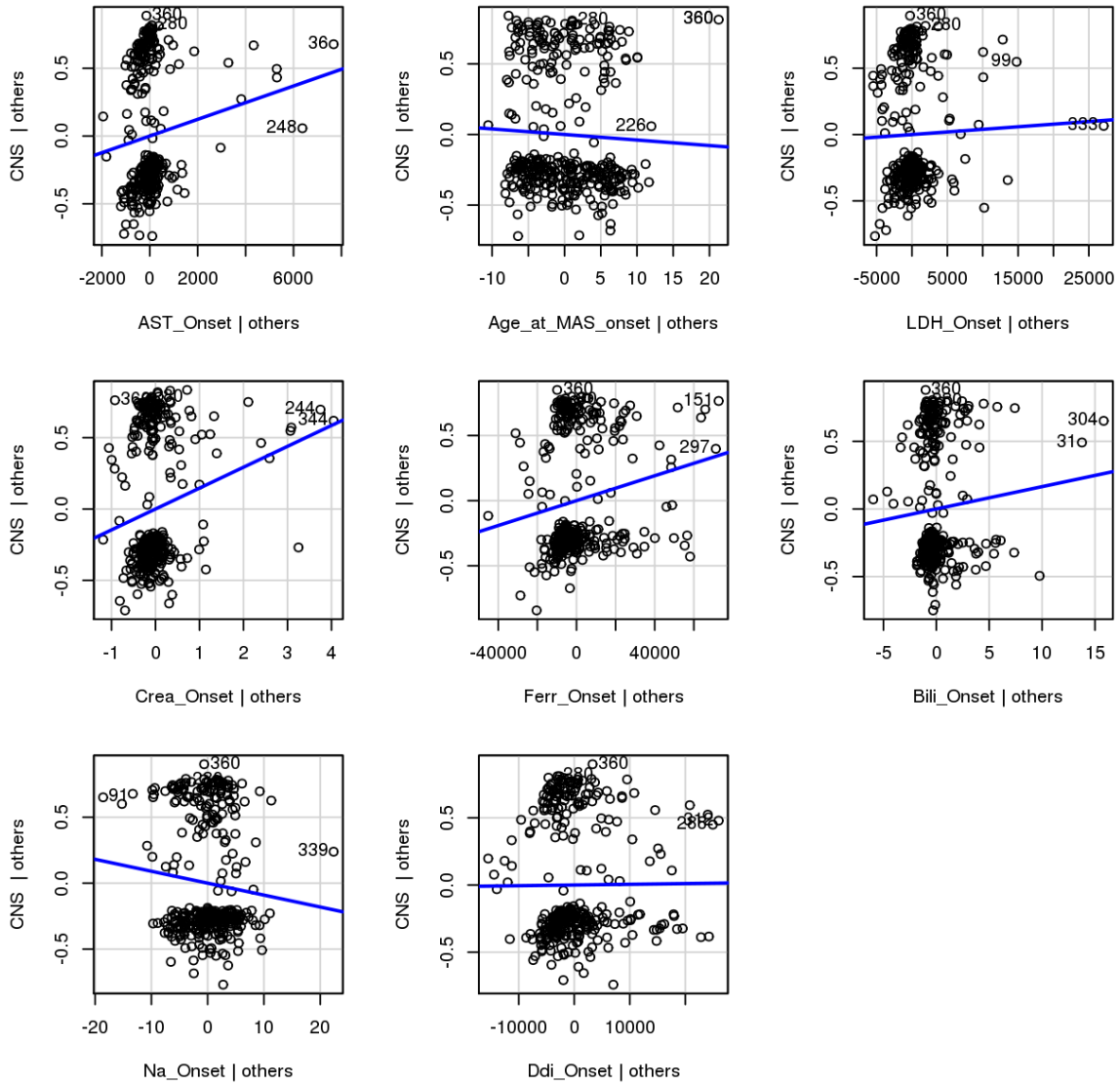


Fig. 33 – Added-variable plots for selected predictors

3.2 Accounting for treatment effect on outcomes: estimation of causal impact of Cyclosporine and Etoposide therapy through Inverse Probability Weighting (IPW)

Immunosuppressive treatments have a major impact on outcomes. Ignoring treatment effects in prognostic models limits the accuracy and generability of results[74]. We therefore investigated the causal effect of these treatments on severe course and mortality. In this cohort, the 97,7% of patients were given corticosteroids, precluding any comparison between treated and not-treated subjects. Cyclosporine was used in 61.2 %, while Etoposide in 40%. These treatments were plausibly administered in most severe cases[75], with difference according to the geographic and specialty settings. To accounting for this unbalance in treatment effect estimations, we employed augmented IPW with double robust methods [76]. The adjustment sets were based on the DAGs for severe course and death, shown in Fig 34. We found no evidence for a significant effect of cyclosporine or etoposide on severe course and death rate. Tab. 32 shows the Potential-outcome means (POMs) and Average Treatment Effect for the Treated (ATET) coefficients, p-values and odd ratios for both treatments. Tab. 33 shows the standardized differences between the etoposide treated and not-treated groups before and after IPW for all the included parameters.

[74] Pajouheshnia, R., Damen, J.A.A.G., Groenwold, R.H.H. et al. Treatment use in prognostic model research: a systematic review of cardiovascular prognostic studies. Diagn Progn Res 1, 15 (2017) doi:10.1186/s41512-017-0015-0

[75] Minoia, Francesca, et al. "Dissecting the heterogeneity of macrophage activation syndrome complicating systemic juvenile idiopathic arthritis." The Journal of rheumatology 42.6 (2015): 994-1001.

[76] Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. Statistics in Medicine. 2015; 34(28):3661–3679.

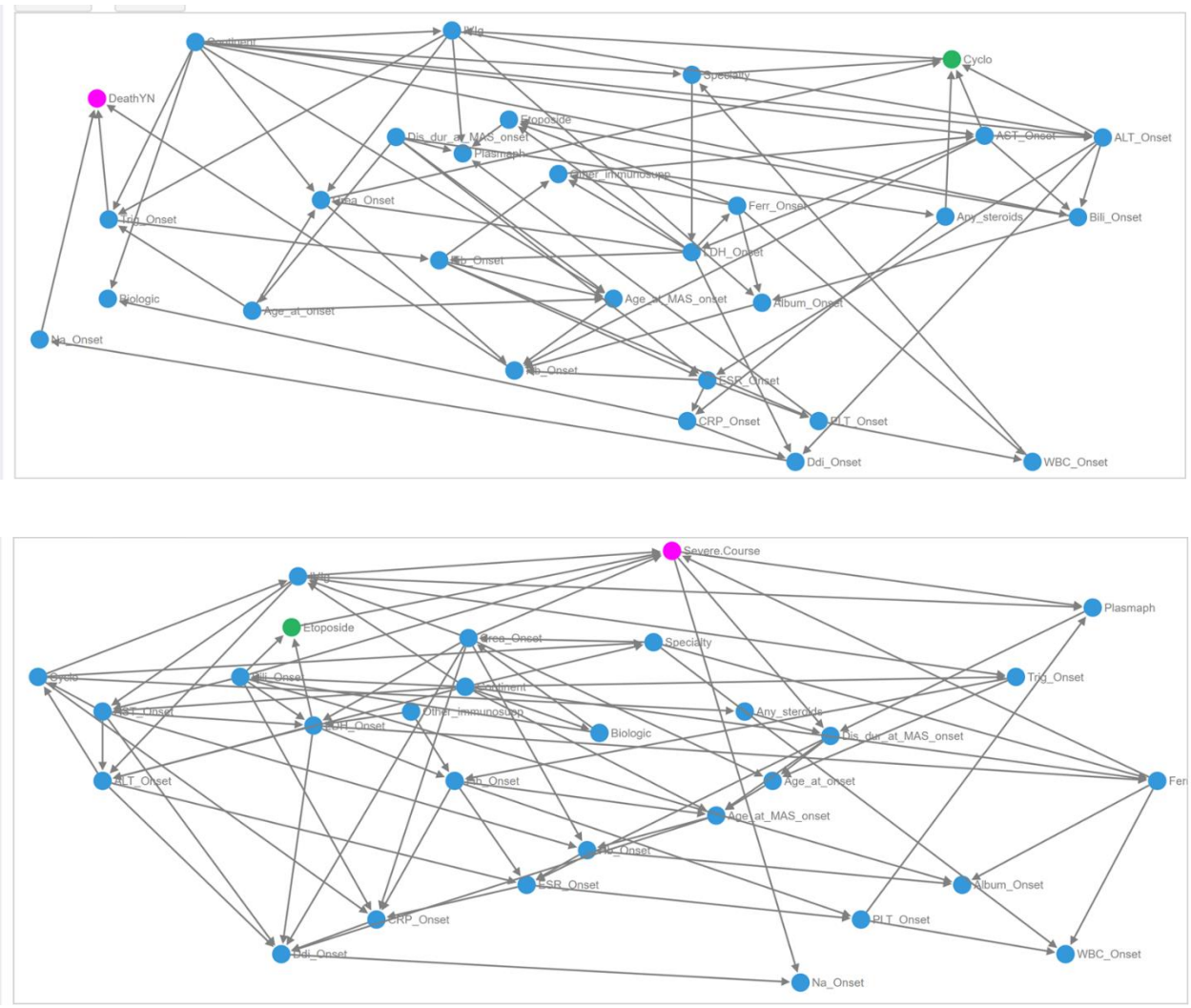


Fig. 34 – DAGs for the death and severe course risk including treatments and other potential counfounders.

```
Treatment-effects estimation                    Number of obs     =        335
Estimator      : IPW regression adjustment
Outcome model  : logit
Treatment model: logit
```

SevereCourse	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
ATET						
Cyclo (1 vs 0)	**.0469724**	**.0649138**	**0.72**	**0.469**	**-.0802563**	**.1742011**
POmean						
Cyclo 0	**.2649088**	**.0575352**	**4.60**	**0.000**	**.1521419**	**.3776757**

```
Iteration 0:   EE criterion =  3.277e-16
Iteration 1:   EE criterion =  9.606e-30

Treatment-effects estimation                    Number of obs     =        329
Estimator      : IPW regression adjustment
Outcome model  : logit
Treatment model: logit
```

SevereCourse	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
ATET						
Etoposide (1 vs 0)	**.1206175**	**.0952545**	**1.27**	**0.205**	**-.066078**	**.3073129**
POmean						
Etoposide 0	**.3793825**	**.0679545**	**5.58**	**0.000**	**.2461942**	**.5125708**

```
Iteration 0:   EE criterion =  2.506e-16
Iteration 1:   EE criterion =  5.078e-30

Treatment-effects estimation                    Number of obs     =        339
Estimator      : IPW regression adjustment
Outcome model  : linear
Treatment model: logit
```

DeathYN	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
ATET						
Etoposide (1 vs 0)	**.0312645**	**.0633927**	**0.49**	**0.622**	**-.0929829**	**.1555118**
POmean						
Etoposide 0	**.1187355**	**.0375747**	**3.16**	**0.002**	**.0450905**	**.1923806**

Tab. 32 – Potential-outcome means (POMs) and Average Treatment Effect for the Treated (ATET) coefficients, p-values and odd ratios obteined by means of IPW for both treatments

POmean						
Etoposide						
0	.3793825	.0679545	5.58	0.000	.2461942	.5125708

. . tebalance summarize

Covariate balance summary

	Raw	Weighted
Number of obs =	329	329.0
Treated obs =	38	167.7
Control obs =	291	161.3

	Standardized differences		Variance ratio	
	Raw	Weighted	Raw	Weighted
Bili_Onset	.4219413	.1201211	5.206968	1.961784
Crea_Onset	.3412794	.0004783	3.572741	1.228052
Ferr_Onset	.3596005	.1011162	2.708151	1.473961
IVIg	.2794255	.0423385	1.142417	1.006064
Hb_Onset	-.4584765	.0003631	.7547063	.6343863
Trig_Onset	.425364	-.0944713	.7235454	.2017616
Ddi_Onset	.4793018	.0957291	3.15244	1.619714
Specialty	-.2972076	-.145118	1.390373	1.128482

Tab. 33 – Standardized differences between the etoposide treated and not-treated groups before and after IPW for the included parameters.

4. Discussion

In this work we aimed to apply a causal inference approach to the evaluation of prognostic biomarkers in sJIA-MAS. In order to investigate the causal pathways leading to unfavorable outcomes in these patients, we used theory and data-based graphical models that informed mediation analyses and variable selection for unbiased estimation of direct and indirect effect of measurable alterations at disease onset.

This allows to identify which, among factors that have been previously reported to be associated with increased risk, are causal determinants of outcomes and therefore possible targets of therapeutic interventions.

Moreover, the causal analysis framework allowed for the exploration of specific treatment which, in absence of clinical trials, have not been empirically evaluated before. Directed acyclic graphs are valuable algebraic tools to distinguish causal from spurious associations by incorporating expert knowledge in the model. DAGs design require the a priori specification of plausible causal effects between variables of interest; although this introducing an inherent element of subjectivity in the analysis, it allows to expose transparently all the assumptions underlying the model.

DAG-informed regression analysis demonstrated a causal effect of creatinine (AKI) and sodium levels on mortality, confirming our previous findings. Hyponatremia emerged as an independent risk factor, not explained by concomitant inflammation-driven Syndrome of inappropriate antidiuretic hormone secretion (SIADH) and AKI. We couldn't directly assess a possible association with hypovolemic or distributive hemodynamic dysfunction (shock); however, mediation analysis evidenced no indirect effect due to heart failure (not shown) or multiple organ failure, which are frequent in similar clinical pictures.

Despite not being identified as independent predictors, other biomarkers, such as AST and LDH, showed a non-linear association with death risk after adjustment with confounding variables; thus, although targeting this particular manifestations would not plausibly reduce mortality, they could provide useful prognostic information when considering specific cut-offs.

Acute renal injury, as measured by creatinine levels, appeared the strongest determinant of death risk. This is in line with previous report of AKI as a predictor of poor outcome in HLH patients[77]. Many mechanisms have been proposed to explain the occurrence of kidney involvement in these forms, including histiocytic glomerular infiltration, direct cytokine-induced damage on glomerular epithelium, and coagulopathy-related disorders[78]. Microvascular thromboses of renal tissue[79] have been demonstrated in patients with disseminated intravascular coagulation who develop AKI.

To unravel the pathways leading to AKI and its association with mortality in MAS patients, we compared two possible pathophysiological models, one considering disseminate intravascular coagulation-like mechanism, the other involving thrombotic microangiopathy, by means of mediation analysis. Known differences in laboratory patterns associated with these two conditions were used to assess and compare the two explanatory models on data. We found that the TMA model, involving a pathway linking platelet decreasing with anemia, high LDH and bilirubin values due to supposed hemolysis, and renal injury was not supported by regression results. Conversely, consumptive coagulopathy (increased d-dimer levels) exhibit an indirect effect on death risk through the association with kidney injury (high creatine); this suggest a causal role for DIC-like coagulative dysfunction in the development of life-threatening kidney involvement in these patients. Further investigations to clarify the mechanism behind kidney

[77] Santoriello, Dominick, Jonathan Hogan, and Vivette D. D'Agati. "Hemophagocytic syndrome with histiocytic glomerulopathy and intraglomerular hemophagocytosis." American Journal of Kidney Diseases 67.6 (2016): 978-983.
[78] Karras, Alexandre. "What nephrologists need to know about hemophagocytic syndrome." Nature Reviews Nephrology 5.6 (2009): 329.
[79] Esmaili, Haydarali, et al. "An update on renal involvement in hemophagocytic syndrome (macrophage activation syndrome)." Journal of nephropathology 5.1 (2016): 8.

injury and other severe manifestations in MAS are needed to guide clinical decision-making in these patients, as the therapeutic approach is different in TMA[80] and DIC-like abnormalities or full-blown DIC[81].

We found evidence that coagulopathy, as measured by fibrinogen levels, is also related to the development of CNS disease. This effect seems only partially (25%) mediated by the occurrence of clinical hemorrhagic manifestations. This complication appears more frequent in patients that were younger al MAS onset, although age was not confirmed as an independent predictor of CNS disease in this analysis. Interestingly, age at MAS onset also moderated the association between ferritin and CNS involvement, with younger patients already showing neurologic manifestations at lower ferritin values.

The strongest determinant of CNS disease was creatinine, showing an OR of 2.4 (95%Cl 1.5 - 3.9). Around 18% of this effect is explained by the occurrence of AKI int the context of a more severe multisystemic dysfunction, as indicated by evidence of a mediation effect by heart failure.

Similarly, creatinine was the main determinant of severe course (OR 1.8, 95%Cl 1.2 - 2.7). A non-linear behavior is evident in this association, with higher probability of occurrence for levels above 0.7 mg/dl.

[80] Scully M, Hunt BJ, Benjamin S, Liesner R, Rose P, Peyvandi F, Cheung B, Machin SJ. British Committee for Standards in Haematology: guidelines on the diagnosis and management of thrombotic thrombocytopenic purpura and other thrombotic microangiopathies. Br J Haematol. 2012;158:323–35.

[81] Taylor FB Jr, Toh CH, Hoots WK, Wada H, Levi M. Scientific subcommittee on disseminated intravascular coagulation (DIC) of the international society on thrombosis and Haemostasis (ISTH): towards definition, clinical and laboratory criteria, and a scoring system for disseminated intravascular coagulation. Thromb Haemost. 2001;86:1327–30

Hypofibrinogenemia emerged as another significant risk factor for severe course, underlining again the causal role of coagulative dysfunction in the development of unfavorable outcomes.

Others parameter that didn't confer a significant risk according to logistic regression analysis, such as ferritin, showed nevertheless a non-linear association with the outcome in MARS model; a threshold of 20000 ng/dl identified a clear increment in the risk of severe outcome.

Although ferritin is generally considered a surrogate marker of histiocytic activation[82], reflecting both IL-18 and IFNγ production, some authors have advanced the hypothesis of a mechanistic role of this molecule in feedforwarding the hyperinflammatory process in MAS and MAS-predisposing conditions[83].

Consistently with a previous analysis[84], non-linear models revealed also a correlation between higher age at SJIA onset and disease duration at MAS diagnosis and occurrence of a severe course. Patients above 11 years of age and with sJIA for more than 3 years appeared at higher risk; moreover, age at MAS onset, in particular for subject above 6 years old, correlated with the unfavorable outcome. However, these demographic features did not result independent predictors in adjusted regression models.

Unsurprisingly, ESR, which resulted an independent predictor in our previous predictive analysis, did not show a significant association with severe course once adjusted for coagulation markers,

[82] Finch CA, Huebers HA, Cazzola M, Bergamaschi G, Bellotti V. Storage iron. In: Albertini A, Arosio P, Chiancone E, Drysdale J, editors. Ferritins and isoferritins as biochemical markers. Amsterdam, The Netherlands: Elsevier Science Publishers; 1984.
[83] Ruscitti, Piero, et al. "Prognostic factors of macrophage activation syndrome, at the time of diagnosis, in adult patients affected by autoimmune disease: analysis of 41 cases collected in 2 rheumatologic centers." Autoimmunity reviews 16.1 (2017): 16-21.
[84] Minoia, Francesca, et al. "Dissecting the heterogeneity of macrophage activation syndrome complicating systemic juvenile idiopathic arthritis." The Journal of rheumatology 42.6 (2015): 994-1001.

confirming that the prognostic information conveyed by ESR just reflects the level of fibrinogen consumption. Interestingly, no effect on ESR values nor impact on outcomes were observed for Hb levels, conversely to what emerged from the previous analysis[85].

In the final part of the analysis, we addressed the causal effect of treatments on outcomes. Given the absence of validated evidence-based treatment guidelines on sJIA-MAS, treatment relies on expert knowledge and anecdotical reports[86]. Current management in pediatric rheumatology settings usually involve the use of corticosteroids and cyclosporine; increasingly, biologic agents such as IL-1 blockers are employed[87]. While not considered a primary approach in MAS, etoposide, a cardinal treatment for genetic forms of HLH, remains a therapeutic option for most severe cases[88].

Despite the non-randomized nature of our data, insights on treatments effects within the causal inference framework can be provided using inverse probability weighting[89]. By use of augmented IPW we obtained balanced samples of patients across treatment groups and found no evidence of a significant impact of etoposide and cyclosporine therapies on mortality.

[85] Minoia, Francesca, et al. "Dissecting the heterogeneity of macrophage activation syndrome complicating systemic juvenile idiopathic arthritis." The Journal of rheumatology 42.6 (2015): 994-1001.
[86] Boom, V., et al. "Evidence-based diagnosis and treatment of macrophage activation syndrome in systemic juvenile idiopathic arthritis." Pediatric Rheumatology 13.1 (2015): 55.
[87] Behrens, E. M., & Koretzky, G. A. (2017). Cytokine storm syndrome: Looking toward the precision medicine era. Arthritis & Rheumatology, 69(6), 1135-1143.
[88] Ehl, S., Astigarraga, I., von Bahr Greenwood, T., Hines, M., Horne, A., Ishii, E., ... & Machowicz, R. (2018). Recommendations for the use of etoposide-based therapy and bone marrow transplantation for the treatment of HLH: consensus statements by the HLH Steering Committee of the Histiocyte Society. The Journal of Allergy and Clinical Immunology: In Practice.
[89] Halpern, Elkan F. "Behind the numbers: inverse probability weighting." Radiology 271.3 (2014): 625-628.

However, we couldn't account for differences in the timing of treatment initiation, as temporal information was not included in the study survey.

Also, due to the small number of treated patients in the dataset, which was collected between 2002 and 2012, we couldn't examine the impact of now widespread biologic therapies.

Several other limitations must be acknowledged in the study, not only due to the retrospective and not-randomized nature of the dataset, but also regarding the absence of many not-laboratoristic data, such as vital parameters, that could better explain and predict severe outcomes in acute settings. In particular, the analysis did not consider in the model important underlying pathogenetic factors, such as molecular biomarkers and cytokines levels, that are known to drive biomarkers abnormalities and clinical manifestations. A growing body of evidence demonstrated the key role of specific pathways, driven by IL-18 and IFNγ whose dysregulation strictly correlates with markers of hyperinflammation and organ involvement, including ferritin levels, white blood cell and platelet counts, transaminases and LDH.[90] [91]

However, as cytokine dosing were not object of the original study, we explored only the relationship between routine avaible laboratory parameters, which are plausibly influenced by multiple pathological processes. These unmeasured variables could have affected the nature of the DAGs and thus the results. The integration of organ damage markers and the cytokine networks underlying the immunopathological mechanism of MAS is warranted in further

[90] Shimizu, M., Yokoyama, T., Yamada, K., Kaneda, H., Wada, H., Wada, T., ... & Yachie, A. (2010). Distinct cytokine profiles of systemic-onset juvenile idiopathic arthritis-associated macrophage activation syndrome with particular emphasis on the role of interleukin-18 in its pathogenesis. Rheumatology, 49(9), 1645-1653.
[91] Bracaglia, Claudia, Giusi Prencipe, and Fabrizio De Benedetti. "Macrophage activation syndrome: different mechanisms leading to a one clinical syndrome." Pediatric Rheumatology 15.1 (2017): 5.

research to develop an accurate, multilayered model able to inform a rational choice and prioritization of interventions, in order to improve the outcomes of this severe condition.

Despite these limitations, these findings could have both clinical and research implications. The identifications of determinants of an unfavorable course could guide patient stratification and clinical decision making, increasing the awareness of relevant factors that clinicians need to target therapeutically to improve the outcome. From a methodologic point of view, our work highlights the limits of traditional analytic techniques in prognostic research where a complex interplay of variables is involved, and the potential of a causal inference approach in unraveling this network of effects.

Conclusions

Using a causal inference approach to explore the relationships between clinical features and outcomes in patients with sJIA-MAS, serum creatinine (OR 1.68, 95%Cl 1.11 – 2.50) and sodium values (OR 0.89, 95%Cl 0.82 – 0.96) emerged as main determinants of mortality. Mediation models supported a causal pathway linking coagulopathy to kidney injury and consequentially raised risk of death. Creatinine levels (OR 1.88, 95%Cl 1.29 – 2.76) were also the strongest risk factor for a severe course, followed by hypofibrinogenemia (OR 0.64, 95%Cl 0.48 – 0.84) and LDH (OR 1.34, 95%Cl 1.06 – 1.70). CNS disease was predicted by creatinine (OR 2.4, 95%Cl 1.5 – 3.9) and fibrinogen (OR 0.99, 95%Cl 0.99 -0.99) levels; ferritin levels (OR 1.59, 95%Cl 1.33 -2.15) also showed an effect on the risk of CNS involvement, which was moderated by age at MAS onset.

Acknowledgments

I would like to thank Dr. Louis Anthony (Tony) Cox for giving me the opportunity to use the Causal Analytics Toolkit (CAT) software for this analysis (https://regulatorystudies.columbian.gwu.edu/causal-analytics-toolkit-cat-assessing-potential-causal-relations-data)

References

- Arnold, Kellyn F., et al. "DAG-informed regression modelling, agent-based modelling and microsimulation modelling: a critical comparison of methods for causal inference." International journal of epidemiology 48.1 (2018): 243-253.
- Arnold, Kellyn F., et al. "Generalised linear models for prognosis and intervention: Theory, practice, and implications for machine learning." arXiv preprint arXiv:1906.01461 (2019).
- Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. Statistics in Medicine. 2015; 34(28):3661–3679.
- Behrens, E. M., & Koretzky, G. A. (2017). Cytokine storm syndrome: Looking toward the precision medicine era. Arthritis & Rheumatology, 69(6), 1135-1143.
- Boom, V., et al. "Evidence-based diagnosis and treatment of macrophage activation syndrome in systemic juvenile idiopathic arthritis." Pediatric Rheumatology 13.1 (2015): 55.
- Bracaglia, Claudia, Giusi Prencipe, and Fabrizio De Benedetti. "Macrophage activation syndrome: different mechanisms leading to a one clinical syndrome." Pediatric Rheumatology 15.1 (2017): 5.
- Cattaneo, Matias D., David M. Drukker, and Ashley D. Holland. "Estimation of multivalued treatment effects under conditional independence." The Stata Journal 13.3 (2013): 407-450.
- Crayne, Courtney B., et al. "The Immunology of Macrophage Activation Syndrome." Frontiers in immunology 10 (2019).
- Davì, Sergio, et al. "Macrophage Activation Syndrome." Pediatric Rheumatology. Springer, Singapore, 2017. 275-292.
- E. Vittinghoff et al., Regression Methods in Biostatistics, Statistics for Biology
- Ehl, S., Astigarraga, I., von Bahr Greenwood, T., Hines, M., Horne, A., Ishii, E., ... & Machowicz, R. (2018). Recommendations for the use of etoposide-based therapy and bone marrow transplantation for the treatment of HLH: consensus statements by the HLH Steering Committee of the Histiocyte Society. The Journal of Allergy and Clinical Immunology: In Practice.
- Esmaili, Haydarali, et al. "An update on renal involvement in hemophagocytic syndrome (macrophage activation syndrome)." Journal of nephropathology 5.1 (2016): 8.
- Ferguson, Karl D., et al. "Evidence synthesis for constructing directed acyclic graphs (ESC-DAGs): a novel and systematic method for building directed acyclic graphs." International journal of epidemiology (2019).

- Finch CA, Huebers HA, Cazzola M, Bergamaschi G, Bellotti V. Storage iron. In: Albertini A, Arosio P, Chiancone E, Drysdale J, editors. Ferritins and isoferritins as biochemical markers. Amsterdam, The Netherlands: Elsevier Science Publishers; 1984.
- Friedman, Jerome H., and Charles B. Roosen. "An introduction to multivariate adaptive regression splines." (1995): 197-217.
- Gagliardi, Luigi. "Prediction and causal inference." Acta Paediatrica 98.12 (2009): 1890-1892.
- Greenwell, Brandon M. "pdp: an R Package for constructing partial dependence plots." The R Journal 9.1 (2017): 421-436.
- Halpern, Elkan F. "Behind the numbers: inverse probability weighting." Radiology 271.3 (2014): 625-628.
- Karras, Alexandre. "What nephrologists need to know about hemophagocytic syndrome." Nature Reviews Nephrology 5.6 (2009): 329.
- Karras, Alexandre. "What nephrologists need to know about hemophagocytic syndrome." Nature Reviews Nephrology 5.6 (2009): 329.
- Leopold SJ, Watson JA, Jeeyapant A, Simpson JA, Phu NH, Hien TT, et al. (2019) Investigating causal pathways in severe falciparum malaria: A pooled retrospective analysis of clinical studies. PLoS Med 16(8): e1002858
- Minoia F, Tibaldi J, Muratore V, et al. FRI0565 A MULTINATIONAL STUDY OF THROMBOTIC MICROANGIOPATHY IN MACROPHAGE ACTIVATION SYNDROME: A DREADFUL CONDITION WHICH IS LIKELY UNDER-RECOGNIZEDAnnals of the Rheumatic Diseases 2019;78:978.
- Minoia, Francesca, et al. "Dissecting the heterogeneity of macrophage activation syndrome complicating systemic juvenile idiopathic arthritis." The Journal of rheumatology 42.6 (2015): 994-1001.
- Minoia, Francesca, et al. "Dissecting the heterogeneity of macrophage activation syndrome complicating systemic juvenile idiopathic arthritis." The Journal of rheumatology 42.6 (2015): 994-1001.
- Minoia, Francesca, et al. "Dissecting the heterogeneity of macrophage activation syndrome complicating systemic juvenile idiopathic arthritis." The Journal of rheumatology 42.6 (2015): 994-1001.
- Morgan SL, Winship C, Counterfactuals and Causal Inference: Methods and Principles for Social Research. New York: Cambridge University Press, 2007 of Educational Psychology. 1974; 66(5):688
- Pajouheshnia, R., Damen, J.A.A.G., Groenwold, R.H.H. et al. Treatment use in prognostic model research: a systematic review of cardiovascular prognostic studies. Diagn Progn Res 1, 15 (2017) doi:10.1186/s41512-017-0015-0
- Pearl J. Causality: Models, Reasoning and Inference. Cambridge, UK: Cambridge University Press;
- Ravelli, A., et al. "Macrophage Activation Syndrome." Handbook of Systemic Autoimmune Diseases. Vol. 11. Elsevier, 2016. 85-106.
- Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies. Journal

- Ruscitti, Piero, et al. "Prognostic factors of macrophage activation syndrome, at the time of diagnosis, in adult patients affected by autoimmune disease: analysis of 41 cases collected in 2 rheumatologic centers." Autoimmunity reviews 16.1 (2017): 16-21
- Santoriello, Dominick, Jonathan Hogan, and Vivette D. D'Agati. "Hemophagocytic syndrome with histiocytic glomerulopathy and intraglomerular hemophagocytosis." American Journal of Kidney Diseases 67.6 (2016): 978-983.
- Schooling, C. Mary, and Heidi E. Jones. "Clarifying questions about "risk factors": predictors versus explanation." Emerging themes in epidemiology 15.1 (2018): 10.
- Scully M, Hunt BJ, Benjamin S, Liesner R, Rose P, Peyvandi F, Cheung B, Machin SJ. British Committee for Standards in Haematology: guidelines on the diagnosis and management of thrombotic thrombocytopenic purpura
- Shimizu, M., Yokoyama, T., Yamada, K., Kaneda, H., Wada, H., Wada, T., ... & Yachie, A. (2010). Distinct cytokine profiles of systemic-onset juvenile idiopathic arthritis-associated macrophage activation syndrome with particular emphasis on the role of interleukin-18 in its pathogenesis. Rheumatology, 49(9), 1645-1653.
- Taylor FB Jr, Toh CH, Hoots WK, Wada H, Levi M. Scientific subcommittee on disseminated intravascular coagulation (DIC) of the international society on thrombosis and Haemostasis (ISTH): towards definition, clinical and laboratory criteria, and a scoring system for disseminated intravascular coagulation. Thromb Haemost. 2001;86:1327–30

Chapter 5 – Development of a probabilistic prognostic model and score for sJIA-MAS

1. Introduction

Macrophage activation syndrome is a severe complication of systemic juvenile idiopathic arthritis (sJIA), belonging to the spectrum of hemophagocytic syndromes (HLH) characterized by a systemic hyperinflammatory process that underlies a wide range of clinical and laboratory alterations, potentially complicating with multiple organ dysfunction and eventually leading to a fatal outcome[92].

Despite the severity of this condition, the contributing factors to the development of outcomes in these patients are not fully understood. Very few evidences are avaible on prediction of outcomes in the HLH spectrum, and no prognostic tool for sJIA-MAS has been developed so far.

The development of multiorgan failure, CNS involvement, anemia and age at onset of MAS older than 11.5 were found to be independent predictors of MAS severity[93]. However, the clinical-laboratory features and organ involvements seen in MAS, including hyperferritinemia, pancytopenia, hepatosplenomegaly, hepatic dysfunction, encephalopathy and coagulation abnormalities, are not independent from one another, but are involved in a complex network of interactions and mutual influences. Previous analyses in the current work found evidences of nonlinearity and interactions between risk factors. Thus, more than the separate impacts of

[92] Ravelli, A., et al. "Macrophage Activation Syndrome." Handbook of Systemic Autoimmune Diseases. Vol. 11. Elsevier, 2016. 85-106.
[93] Minoia, Francesca, et al. "Dissecting the heterogeneity of macrophage activation syndrome complicating systemic juvenile idiopathic arthritis." The Journal of rheumatology 42.6 (2015): 994-1001.

single factors, a more effective modeling strategy could be that of incorporating the relationships among different factors in order to uncover the causal pathways through which the combined effects of multiple predictors determine disease outcome[94].

Bayesian networks (BNs) are a powerful tool that allow to depict the causal relationships between prognostic factors in probabilistic terms, generating quantitative counterfactual predictions for specific clinical scenarios that account for nonlinear relationships[95]. In BNs, the amount of information contributed from prognostic variables can be expressed in information theoretical measures, explicitly modelling the uncertainty in prediction[96].

In this study, we explored the potential of BN and information-theoretical measures to prioritize risk factors and develop a prognostic model for sJIA-MAS.

For this task, we took advantage of the ability of BNs to incorporate prior knowledge in the model, using emerged insights from previous explorative analyses in this work to define the optimal cut-off values for outcomes prediction.

2. Methods

2.1 Patients and data collection

Retrospective patient data were gathered in a multicenter web-based survey involving pediatric rheumatologists and hemato-oncologists from 33 countries. To be included in the study, patients

[94] Thornley S, Marshall RJ, Wells S, Jackson R (2013) Using Directed Acyclic Graphs for Investigating Causal Paths for Cardiovascular Disease. J Biomet Biostat 4: 182. doi:10.4172/2155-6180.1000182

[95] Friedman N, Geiger D, Goldszmidt M (1997) Bayesian network classifiers. Machine Learning 29:131–163.

[96] Si S, Liu G, Cai Z, Xia P (2011) Using Bayesian networks and importance measures to identify tumor markers for breast cancer. In Proceedings of the 2011 IEEE International Conference on Industrial Engineering and Engineering Management, Singapore. Piscataway, IEEE: 1826–1830

had to have SJIA and to have had an episode of MAS diagnosed by the attending physician; treatment had to be specifically altered to treat MAS. Demographic information such as age at MAS onset, age at onset of systemic JIA, duration of systemic JIA at MAS onset, as well as standardized laboratory values collected at disease onset (defined as the time when the initial abnormalities suggesting the occurrence of MAS were detected) were included in the current analysis[97]. A detailed description of the cohort has been previously reported by Minoia et al[98].

2.2 Bayesian Networks

Bayesian networks are a non-parametric formalism which models a network of interacting variables as a graph; the variables are represented as nodes and their interactions as directed links[99]. In other words, the edges in the network encode statistical dependencies between random variables. This representation allows to decompose the joint probability distribution of the variables into marginal probability distributions for the input nodes and conditional probabilities (CP) for other nodes. Thus, BNs can be used not only to map probabilistic dependencies among predictors and outcomes, but allow for manipulating probabilities to draw inferences and answer "what-if" questions, for which the answers are expressed as posterior marginal distributions for the other variables, conditioned on a user-specified set of values[100]. Another advantage of the method is that missing values can be easily handled.

[97] Minoia, Francesca, et al. "Clinical features, treatment, and outcome of macrophage activation syndrome complicating systemic juvenile idiopathic arthritis: a multinational, multicenter study of 362 patients." Arthritis & Rheumatology 66.11 (2014): 3160-3169.
[98] Minoia, Francesca, et al. "Dissecting the heterogeneity of macrophage activation syndrome complicating systemic juvenile idiopathic arthritis." The Journal of rheumatology 42.6 (2015): 994-1001.
[99] Ben-Gal I. Bayesian Networks. In: Ruggeri F, Faltin F, Kenett R, editors. Encyclopedia of Statistics in Quality & Reliability. 2013. Chichester: Wiley and Sons; 2007
[100] Cox Jr, Louis Anthony, Douglas A. Popken, and Richard X. Sun. Causal Analytics for Applied Risk Analysis. Springer International Publishing, 2018.

Among the different algorithm developed to learn BNs from the data, the tree augmented naive Bayes (TAN), which accounts for dependence between variables, is proven to be superior for classification task and therefore has been used successfully in prognostic modelling[101]. This is a supervised data-mining method, where the researcher must select a target variable to be predicted. We accordingly applied a TAN algorithm to learn the structure and network parameters from the dataset for each of the outcome of interest – severe outcome, death and CNS dysfunction, including variables avaible at MAS onset, i.e. demographic features and laboratory values at diagnosis.

As BNs only operates with discrete states, continuous variables were transformed into discrete intervals before computations. For emerged predictors, we used the results from previous analyses in this work to define the optimal intervals. For other parameters, a supervised algorithm was used to identify the most informative cut-offs.

The TAN model was cross-validated 10-fold and the performance assessed through ROC curve analysis.

Once the TAN model was estimated, in order to identify the most important predictors and strongest relationship, we examined the impact of different factors on outcomes distribution using measures drawn upon information theory, such as entropy, mutual information and Kullback-Leibler Divergence[102].

[101] Banu, Bazila, and Ponniah Thirumalaikolundusubramanian. "Comparison of Bayes Classifiers for Breast Cancer Classification." Asian Pacific journal of cancer prevention: APJCP 19.10 (2018): 2917.
[102] Nicholson, Ann E., and Nathalie Jitnah. "Using mutual information to determine relevance in Bayesian networks." Pacific rim international conference on artificial intelligence. Springer, Berlin, Heidelberg, 1998.

Entropy (H) describes the expected amount of information that prognostic variable (X) contains, as defined as:

$$H(X) = -\sum_i P(x_i) \log_2 P(x_i)$$

where P(xi) is the probability of occurrence of X resulting in the possible value xi[103]. Maximizing the entropy equivales to maximizing the information content per observation[104].

Mutual Information (MI) is a measure of how much information is communicated, on average, in one random variable about another; in other words, it measures how much knowing one of these variables reduces our uncertainty about the other[105]. For example, if two variables are independent, knowing one does not provide any information about the other and vice versa, their mutual information will be zero.

For two random variables X and Y, whose joint distribution is defined by P(X, Y), MI is given by[106]

$$I(X;Y) = \sum_{y \in Y} \sum_{x \in X} p(x,y) \log \left(\frac{p(x,y)}{p(x)p(y)} \right)$$

Conditional mutual information (cMI) quantifies the expected value of the mutual information between two variables when a third variable has already been considered. Thus, cMI can be used to avoid information overlapping [107].

A concept strictly related to MI is **Kullback-Leibler Divergence** (KL), or relative entropy, which measures how one probability distribution is different from a second, reference probability distribution[108].For probability distributions P and Q of a discrete random variable, K is given by the the average of the logarithmic difference between the joint probability distributions P(i) and Q(i), where the average is taken using the probabilities P(i)[109]:

[103] Lee, Joon, and David M. Maslove. "Using information theory to identify redundancy in common laboratory tests in the intensive care unit." BMC medical informatics and decision making 15.1 (2015): 59.
[104] Delgado-Bonal, Alfonso, and Javier Martín-Torres. "Human vision is determined based on information theory." Scientific reports 6 (2016): 36038.
[105] Learned-Miller, Erik G. "Entropy and mutual information." Department of Computer Science, University of Massachusetts, Amherst (2013).
[106] Learned-Miller, Erik G. "Entropy and mutual information." Department of Computer Science, University of Massachusetts, Amherst (2013).
[107] Tsagaris, Vassilis, Nikos Fragoulis, and Christos Theoharatos. "Performance evaluation of image fusion methods." Image Fusion (2011): 71-88.
[108] Wikipedia contributors. (2019, December 13). Kullback–Leibler divergence. In Wikipedia, The Free Encyclopedia. Retrieved 14:33, January 8, 2020, from https://en.wikipedia.org/w/index.php?title=Kullback%E2%80%93Leibler_divergence&oldid=930573126
[109] Kullback, S. (1959), Information Theory and Statistics, John Wiley & Sons. Republished by Dover Publications in 1968; reprinted in 1978: ISBN 0-8446-5625-9.

$$D_{KL} = (P \parallel Q) = \sum_i P(i) \log \frac{P(i)}{Q(i)}$$

KL provides information on the differences between the probability distributions represented by the network with a specific link (i.e., a specific relationship between two variables), comparing to the network without this link; thus, it is useful to derive the contribution of each detected effect to the model. These measures, thus, allow to rank prognostic factors by their contribution to risk prediction.

We computed KL and MI metrics for each included variable. A sensitivity analysis was performed to evaluate how parameters vary in the condition that the value of prior probabilities varies.

We then generated a decision tree to represent the hierarchy of the obtained prognostic criteria, based on entropy (or binary entropy) of the outcome node. The algorithm displays the optimal sequence of evidence that should be sought in order to gain the maximum amount of information on the state of the outcome of interest. The tree identifies distinct strata of patients with different probability of outcomes. reduction of entropy (information gain) provided by the variable provided by the combination of evidences.

Based on the trained classifier, we implemented a decision support tool in the form of an online prediction system that provides probabilities of MAS outcomes according to user-provided values of predictors.

Finally, to improve the usability of the results, a simplified numeric risk score for severe outcome and death was developed assigning to each predictor a weight based on a genetic algorithm that

takes in account the effects on outcomes and the evidences maximizing outcomes probabilities[110].

2.3 Statistical analysis

Bayeslab (Bayesia Limited Company, France) was used to develop the BNs. Accuracy of the simplified score was assessed with the Jamovi software.

3. Results

3.1 Severe Outcome Model

3.1.1 Structure and Parameters of the Bayesian Networks

The structure of the Bayesian Network for severe outcome is depicted in Fig. 1. 15 predictors were linked to the outcome node, namely ferritin, LDH, fibrinogen, d-dimer, creatinine, BUN, sodium, neutrophil count, platelet count, ESR, AST, ALT, bilirubin, albumin, SJIA duration at disease onset. The relationship between AST, ALT and bilirubin were incorporated in the model. To assess the possible information provided by the latent clustering procedure reported in Chapter 2, we fitted a second dole including also the membership probability for each identified cluster. In the graph, a priori probabilities for each variable state (i.e., the distribution of values in the whole sample) are reported in each node. The

[110] Conrady, Stefan, and Lionel Jouffe. Bayesian networks and BayesiaLab: A practical introduction for researchers. Bayesia USA, 2015.

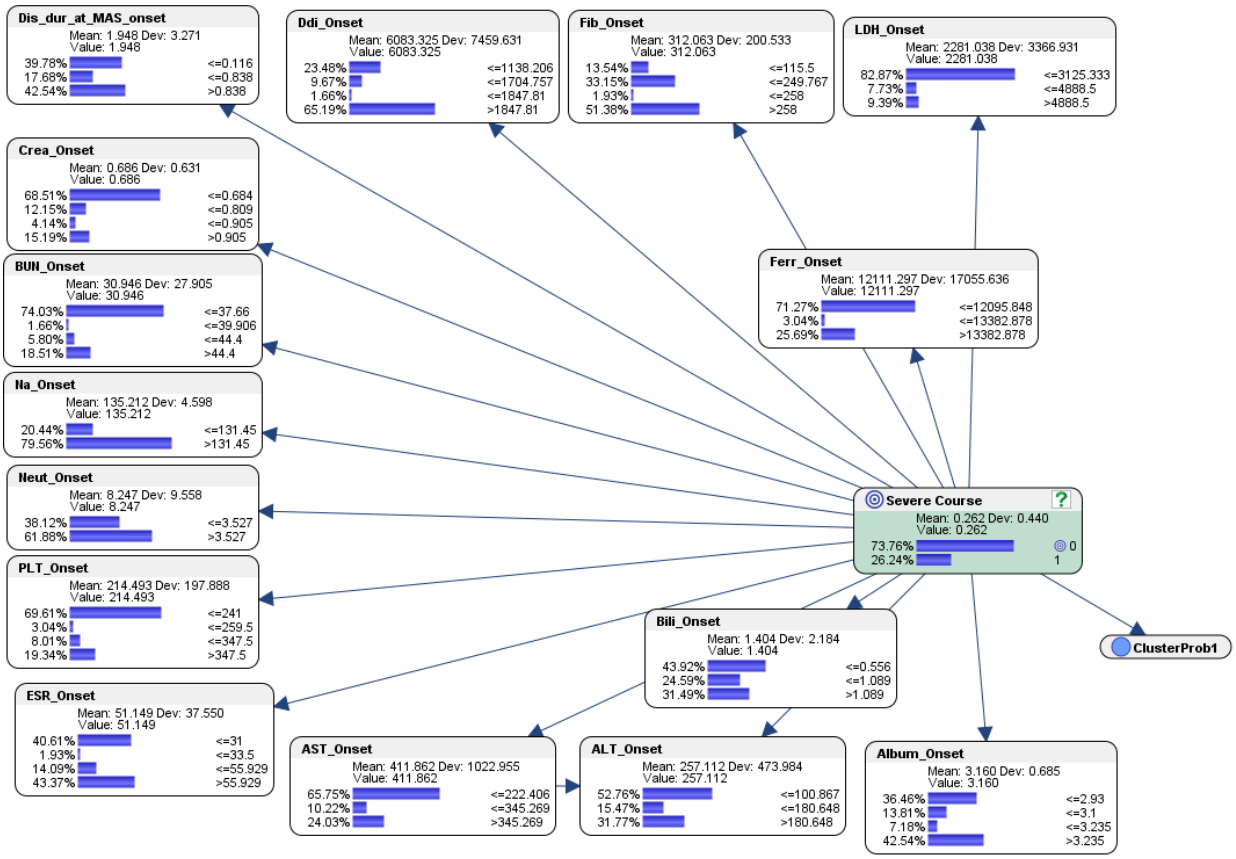


Fig. 1- The structure of the BN for prediction of severe course.

P(Severe Course = 1 | ...)

Fig. 2 -The graphs show the probability of severe course for different values of each parameter.

Histograms in Fig. 2 show the probability of severe course for different value state for each parameter.

3.1.2 Importance ranking of prognostic variables

The importance of prognostic factors was analyzed based on

1) the total and direct effect on the outcome
2) sensitivity analysis
3) Information theoretic metrics

3.1.3 Effects of variables on severe outcome: Which are the causal drivers of severe outcome?

The total effect of each variable, which represents the impact of a small modification of the values of one variable on the outcome, considering the influence of the variable on every other node in the network, is reported in Table 1 and in the quadrant graph below. The positive impacts are displayed in blue and the negative ones are displayed in red. Fibrinogen, sodium e BUN levels emerged as the parameter with the highest impact on outcome.

Total Effects on Target Severe Course									
Node	Prior Value/Mean	Standardized Total Effects	Total Effects	G-test	df	p-value	G-test (Data)	df (Data)	p-value (Data)
Fib_Onset	312.0628	-0.2796	-0.0008	35.4917	3	0.0000%	35.4917	3	0.0000%
Na_Onset	135.2116	-0.2270	-0.0300	17.2129	1	0.0033%	17.2129	1	0.0033%
BUN_Onset	30.9457	0.2236	0.0049	17.3374	3	0.0602%	17.3374	3	0.0602%
Crea_Onset	0.6860	0.2156	0.2019	16.1677	3	0.1048%	16.1677	3	0.1048%
AST_Onset	411.8622	0.2119	0.0002	15.5878	2	0.0412%	15.5878	2	0.0412%
Ferr_Onset	12,111.2970	0.2113	0.0000	15.2945	2	0.0477%	15.2945	2	0.0477%
Album_Onset	3.1603	-0.1904	-0.1490	13.3072	3	0.4017%	13.3072	3	0.4017%
ALT_Onset	257.1122	0.1743	0.0003	10.6533	2	0.4860%	10.6533	2	0.4860%
Ddi_Onset	6,083.3248	0.1726	0.0000	11.4309	3	0.9610%	11.4309	3	0.9610%
ESR_Onset	51.1492	-0.1698	-0.0022	12.5030	3	0.5844%	12.5030	3	0.5844%
LDH_Onset	2,281.0383	0.1505	0.0000	27.9819	2	0.0001%	27.9819	2	0.0001%
Bili_Onset	1.4038	0.1486	0.0474	10.4111	2	0.5486%	10.4111	2	0.5486%
PLT_Onset	214.4928	-0.1471	-0.0004	10.3391	3	1.5893%	10.3391	3	1.5893%
Dis_dur_at_MAS_onset	1.9479	0.1347	0.0280	6.5197	2	3.8394%	6.5197	2	3.8394%
ClusterProb1	0.2448	0.1277	0.1388	5.6655	1	1.7302%	5.6655	1	1.7302%
Neut_Onset	8.2470	0.1062	0.0092	4.1827	1	4.0839%	4.1827	1	4.0839%

Tab. 1 – Total effect of each predictor on severe outcome. Negative relationships are highlighted in red, positive in blue.

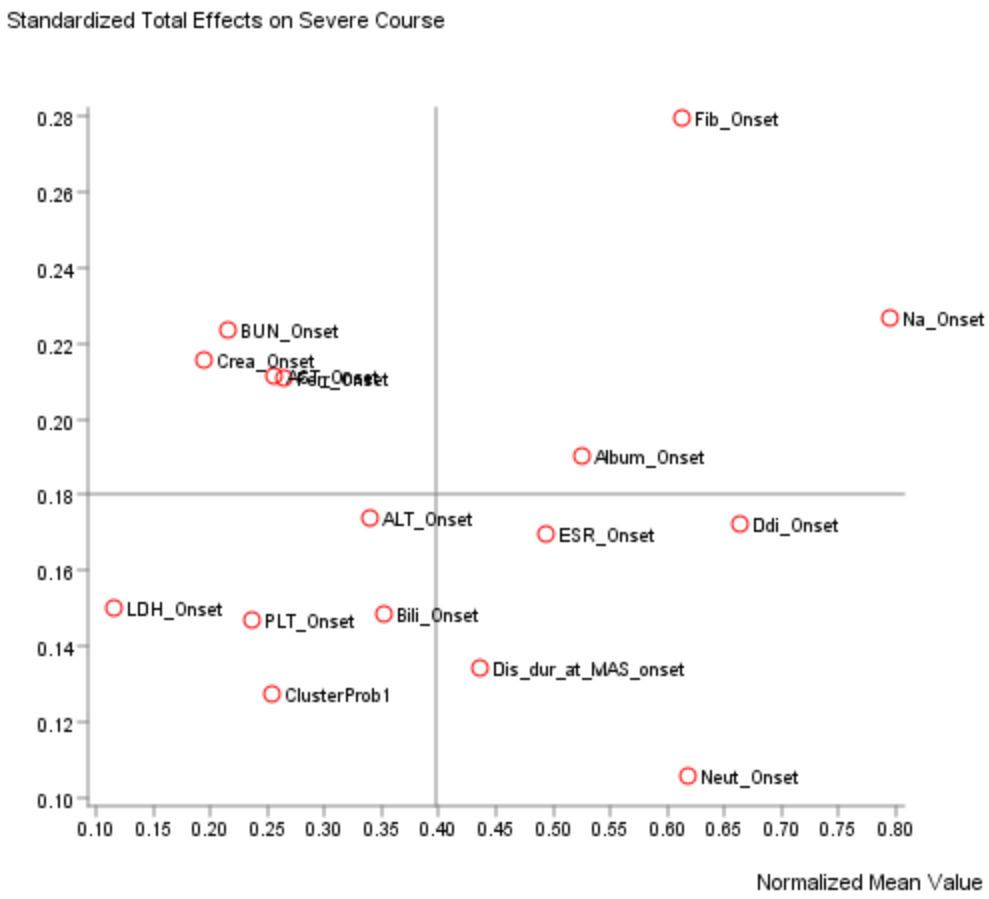


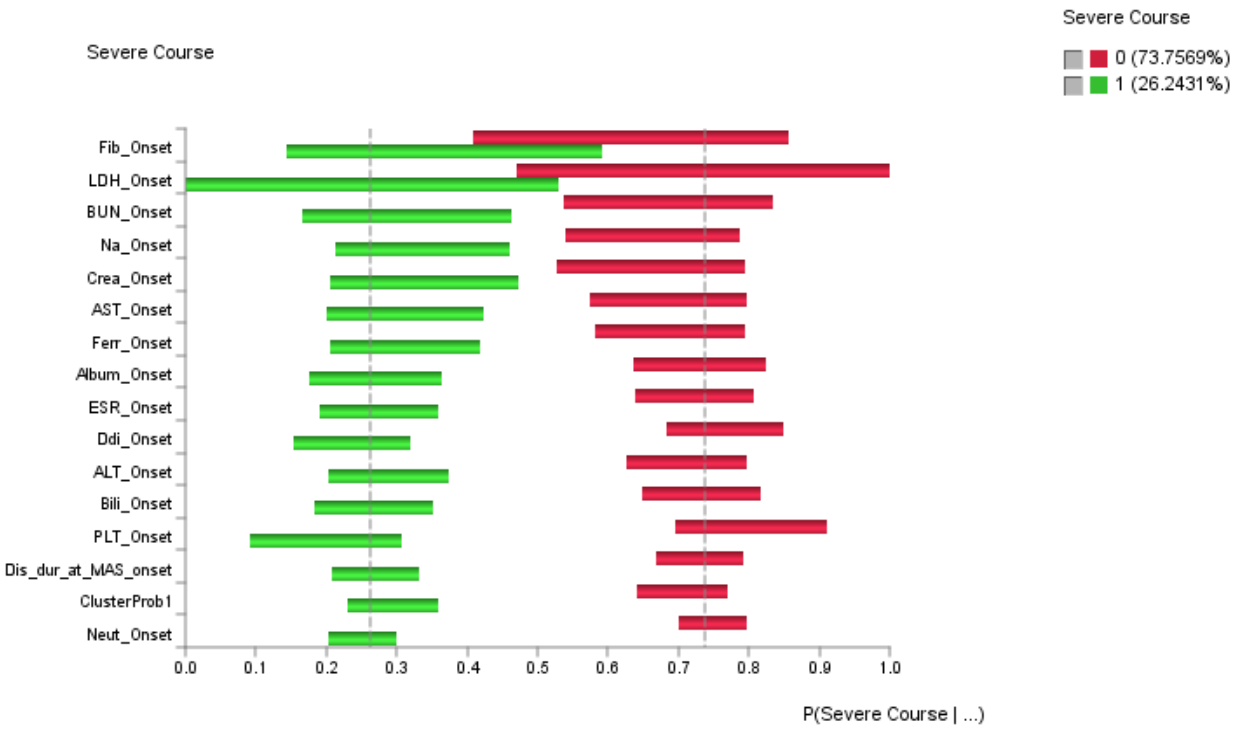
Fig. 3- the Standardized Total Effect displayed in the quadrant graph represents the Total Effect multiplied by the ratio of the standard deviation of the variable and the standard deviation of the outcome

3.1.4 Sensitivity analysis

Sensitivity analysis[111] investigates the effect of small changes in parameter values (i.e., probabilities) on the outcome (e.g., posterior probabilities), computing a set of derivatives of the

[111] Castillo, E., Gutierrez, J. M. & Hadi, A. S. (1997). Sensitivity analysis in discrete Bayesian networks, IEEE Transactions on Systems, Man, and Cybernetics 27(4): 412-423.

posterior probability distributions over each of the value states. Highly sensitive parameters affect the model more significantly. Conversely, If the derivative is small, even large changes in the parameter make little difference in the probability of the outcome[112]. In Fig. 4, sensitivity analyses for severe outcome and for both states of the outcome node are reported, with most sensitive factors, such as LDH, fibrinogen and BUN, represented in the top of the plot.



[112] https://support.bayesfusion.com/docs/GeNIe/bn_sensitivitybn.html

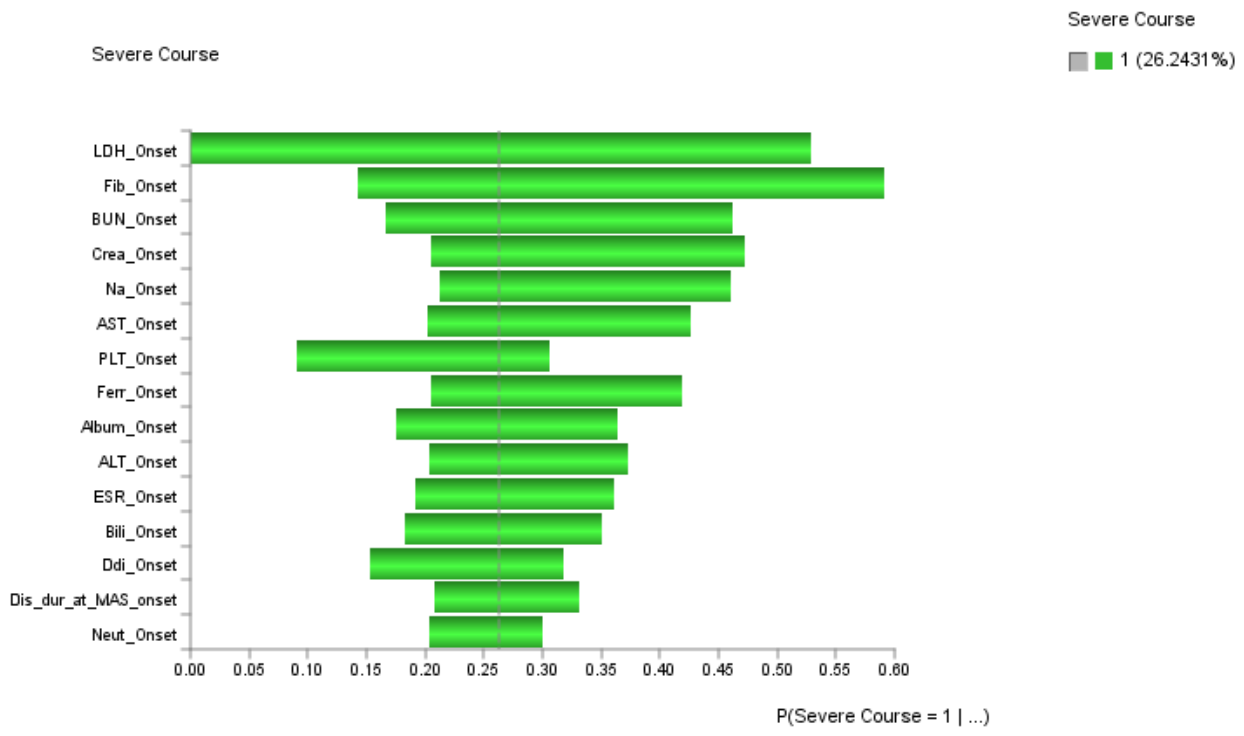


Fig. 4 – Tornado plots depicting the derivatives computed for each variable, expressing the influence of their variance on the model results

3.1.5 Information theoretical metrics: Which are the most informative variables for severe outcome?

A ranking of prognostic factors based on information theoretical indexes, such as MI, is reported in Tab. 2. The cut-off values (variable states) that provide the maximum information gain in forecasting the absence or occurrence of severe outcome are shown in Tab. 3 and 4. The relationship between variable values and their MI is represented in the quadrant plot (Fig. 5). Fibrinogen, LDH, BUN, Na levels exhibited the highest MI. The uncertainty on the outcome lowers with the cumulative evidence provided by these biomarkers, as show in Fig. 6.

Overall Analysis with Severe Course											
Node	Mutual Information	Normalized Mutual Information	Relative Mutual Information	Relative Significance	Prior Mean Value	G-test	df	p-value	G-test (Data)	df (Data)	p-value (Data)
Fib_Onset	0.0707	7.0723%	8.5168%	1.0000	312.0628	35.4917	3	0.0000%	35.4917	3	0.0000%
LDH_Onset	0.0558	5.5759%	6.7147%	0.7884	2,281.0383	27.9819	2	0.0001%	27.9819	2	0.0001%
BUN_Onset	0.0345	3.4548%	4.1604%	0.4885	30.9457	17.3374	3	0.0602%	17.3374	3	0.0602%
Na_Onset	0.0343	3.4300%	4.1305%	0.4850	135.2116	17.2129	1	0.0033%	17.2129	1	0.0033%
Crea_Onset	0.0322	3.2217%	3.8797%	0.4555	0.6860	16.1677	3	0.1048%	16.1677	3	0.1048%
AST_Onset	0.0311	3.1061%	3.7406%	0.4392	411.8622	15.5878	2	0.0412%	15.5878	2	0.0412%
Ferr_Onset	0.0305	3.0477%	3.6702%	0.4309	12,111.2970	15.2945	2	0.0477%	15.2945	2	0.0477%
Album_Onset	0.0265	2.6517%	3.1933%	0.3749	3.1603	13.3072	3	0.4017%	13.3072	3	0.4017%
ESR_Onset	0.0249	2.4914%	3.0003%	0.3523	51.1492	12.5030	3	0.5844%	12.5030	3	0.5844%
Ddi_Onset	0.0228	2.2778%	2.7430%	0.3221	6,083.3248	11.4309	3	0.9610%	11.4309	3	0.9610%
ALT_Onset	0.0212	2.1229%	2.5564%	0.3002	257.1122	10.6533	2	0.4860%	10.6533	2	0.4860%
Bili_Onset	0.0207	2.0746%	2.4983%	0.2933	1.4038	10.4111	2	0.5486%	10.4111	2	0.5486%
PLT_Onset	0.0206	2.0602%	2.4810%	0.2913	214.4928	10.3391	3	1.5893%	10.3391	3	1.5893%
Dis_dur_at_MAS_onset	0.0130	1.2992%	1.5645%	0.1837	1.9479	6.5197	2	3.8394%	6.5197	2	3.8394%
ClusterProb1	0.0113	1.1289%	1.3595%	0.1596	0.2448	5.6655	1	1.7302%	5.6655	1	1.7302%
Neut_Onset	0.0083	0.8335%	1.0037%	0.1179	8.2470	4.1827	1	4.0839%	4.1827	1	4.0839%

Tab 2 – Information theoretical metrics for each predictor

Local Analyzes with Target States										
Severe Course = 0 (73.7569%)										
Node	Binary Mutual Information	Relative Binary Mutual Information	Binary Relative Significance	Posterior Mean Value	Max Bayes Factor			Min Bayes Factor		
Fib_Onset	0.0707	8.5168%	1.0000	338.7735	<=258 (3/4)	2.2472%	1.1621	<=115.5 (1/4)	7.4906%	0.5534
LDH_Onset	0.0558	6.7147%	0.7884	2,041.4156	<=4888.5 (2/3)	10.4869%	1.3558	>4888.5 (3/3)	5.9925%	0.6380
BUN_Onset	0.0345	4.1604%	0.4885	28.2821	<=39.906 (2/4)	1.8727%	1.1298	>44.4 (4/4)	13.4831%	0.7285
Na_Onset	0.0343	4.1305%	0.4850	135.6630	>131.45 (2/2)	85.0187%	1.0686	<=131.45 (1/2)	14.9813%	0.7329
Crea_Onset	0.0322	3.8797%	0.4555	0.6256	<=0.684 (1/4)	73.7828%	1.0770	>0.905 (4/4)	10.8614%	0.7149
AST_Onset	0.0311	3.7406%	0.4392	345.4248	<=222.406 (1/3)	71.1610%	1.0824	>345.269 (3/3)	18.7266%	0.7792
Ferr_Onset	0.0305	3.6702%	0.4309	10,415.5004	<=12095.848 (1/3)	76.7790%	1.0773	>13382.878 (3/3)	20.2247%	0.7872
Album_Onset	0.0265	3.1933%	0.3749	3.2241	>3.235 (4/4)	47.5655%	1.1181	<=2.93 (1/4)	31.4607%	0.8628
ESR_Onset	0.0249	3.0003%	0.3523	54.5605	>55.929 (4/4)	47.5655%	1.0967	<=31 (1/4)	35.2060%	0.8670
Ddi_Onset	0.0228	2.7430%	0.3221	5,690.3680	<=1138.206 (1/4)	26.9663%	1.1484	>1847.81 (4/4)	60.2996%	0.9249
ALT_Onset	0.0212	2.5564%	0.3002	227.4473	<=100.867 (1/3)	56.9288%	1.0790	>180.648 (3/3)	26.9663%	0.8489
Bili_Onset	0.0207	2.4983%	0.2933	1.2812	<=0.556 (1/3)	48.6891%	1.1085	>1.089 (3/3)	27.7154%	0.8801
PLT_Onset	0.0206	2.4810%	0.2913	229.0329	<=259.5 (2/4)	3.7453%	1.2326	<=241 (1/4)	65.5431%	0.9415
Dis_dur_at_MAS_onset	0.0130	1.5645%	0.1837	1.7782	<=0.116 (1/3)	42.6966%	1.0733	>0.838 (3/3)	38.5768%	0.9068
ClusterProb1	0.0113	1.3595%	0.1596	0.2140	<=0.351 (1/2)	77.9026%	1.0445	>0.351 (2/2)	22.0974%	0.8695
Neut_Onset	0.0083	1.0037%	0.1179	7.9255	<=3.527 (1/2)	41.1985%	1.0807	>3.527 (2/2)	58.8015%	0.9503

Severe Course = 1 (26.2431%)										
Node	Binary Mutual Information	Relative Binary Mutual Information	Binary Relative Significance	Posterior Mean Value	Max Bayes Factor			Min Bayes Factor		
Fib_Onset	0.0707	8.5168%	1.0000	236.9917	<=115.5 (1/4)	30.5263%	2.2552	<=258 (3/4)	1.0526%	0.5444
LDH_Onset	0.0558	6.7147%	0.7884	2,954.5042	>4888.5 (3/3)	18.9474%	2.0173	<=4888.5 (2/3)	0.0000%	0.0000
BUN_Onset	0.0345	4.1604%	0.4885	38.4320	>44.4 (4/4)	32.6316%	1.7631	<=39.906 (2/4)	1.0526%	0.6351
Na_Onset	0.0343	4.1305%	0.4850	133.9429	<=131.45 (1/2)	35.7895%	1.7508	>131.45 (2/2)	64.2105%	0.8071
Crea_Onset	0.0322	3.8797%	0.4555	0.8560	>0.905 (4/4)	27.3684%	1.8013	<=0.684 (1/4)	53.6842%	0.7836
AST_Onset	0.0311	3.7406%	0.4392	598.5863	>345.269 (3/3)	38.9474%	1.6206	<=222.406 (1/3)	50.5263%	0.7685
Ferr_Onset	0.0305	3.6702%	0.4309	16,877.3780	>13382.878 (3/3)	41.0526%	1.5980	<=12095.848 (1/3)	55.7895%	0.7828
Album_Onset	0.0265	3.1933%	0.3749	2.9808	<=2.93 (1/4)	50.5263%	1.3856	>3.235 (4/4)	28.4211%	0.6681
ESR_Onset	0.0249	3.0003%	0.3523	41.5615	<=31 (1/4)	55.7895%	1.3739	>55.929 (4/4)	31.5789%	0.7281
Ddi_Onset	0.0228	2.7430%	0.3221	7,187.7403	>1847.81 (4/4)	78.9474%	1.2110	<=1138.206 (1/4)	13.6842%	0.5828
ALT_Onset	0.0212	2.5564%	0.3002	340.4860	>180.648 (3/3)	45.2632%	1.4248	<=100.867 (1/3)	41.0526%	0.7781
Bili_Onset	0.0207	2.4983%	0.2933	1.7484	>1.089 (3/3)	42.1053%	1.3370	<=0.556 (1/3)	30.5263%	0.6950
PLT_Onset	0.0206	2.4810%	0.2913	173.6275	<=241 (1/4)	81.0526%	1.1643	<=259.5 (2/4)	1.0526%	0.3464
Dis_dur_at_MAS_onset	0.0130	1.5645%	0.1837	2.4251	>0.838 (3/3)	53.6842%	1.2619	<=0.116 (1/3)	31.5789%	0.7939
ClusterProb1	0.0113	1.3595%	0.1596	0.3314	>0.351 (2/2)	34.7368%	1.3668	<=0.351 (1/2)	65.2632%	0.8750
Neut_Onset	0.0083	1.0037%	0.1179	9.1506	>3.527 (2/2)	70.5263%	1.1398	<=3.527 (1/2)	29.4737%	0.7732

Tab 3. and 4. – Information theoretical metrics of the predictors for each possible state (absence or occurrence) of severe course.

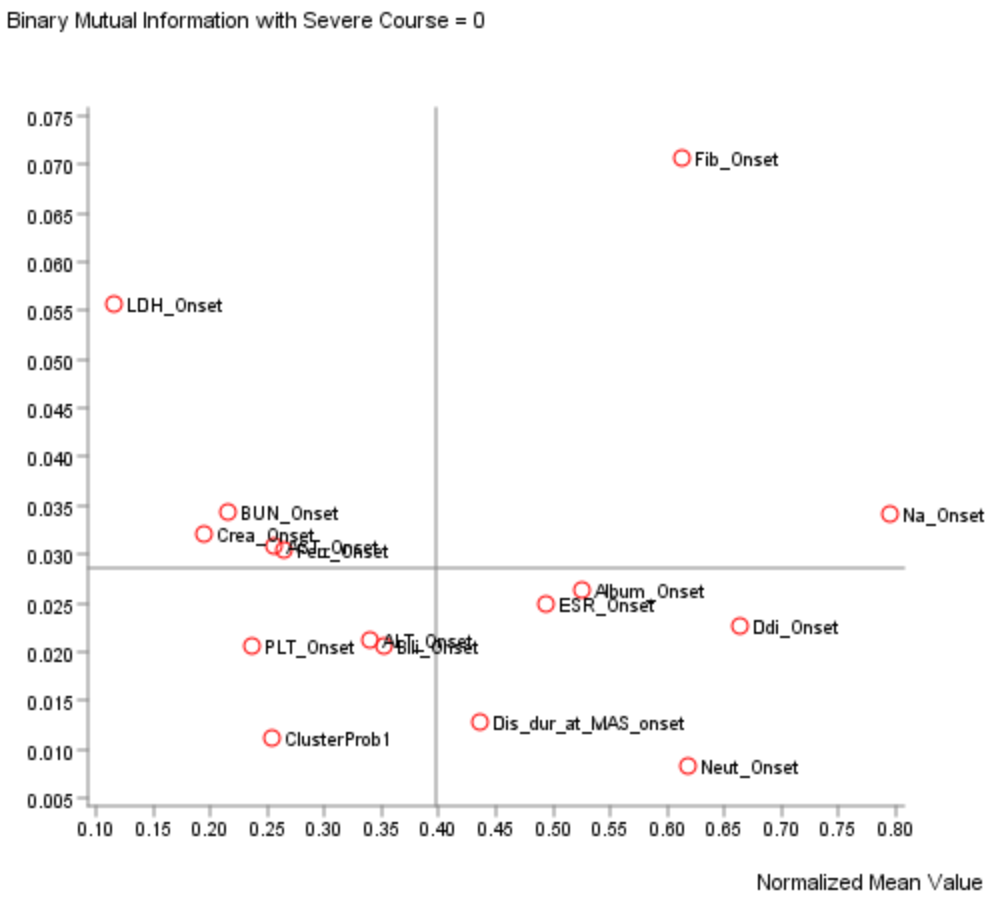


Fig. 5 – The Mutual information of each variable is plotted against its normalized mean value

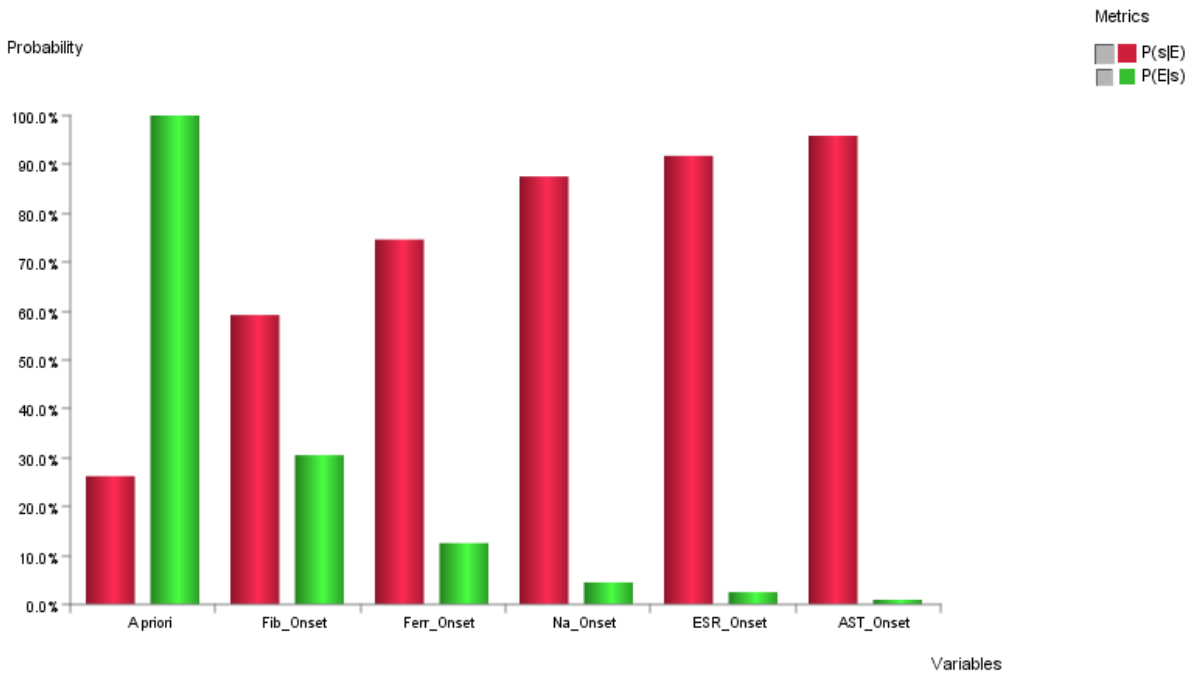


Fig. 6 - The histogram shows how the probability of severe outcome increase by gathering evidence form the selected prognostic biomarkers.

3.1.6 Target Dynamic Profile: Which combination of values are associated with the worst outcome?

Tab. 5 and 6 report the combination of evidences that maximizes the probability of absence and occurrence of severe outcome, respectively.

Dynamic Profile Severe Course: Probability Maximization (A posteriori)						
Search Method: Hard Evidence						
Severe Course = 0						
Node	**Evidence**	**Posterior Probability P(s|E)**	**Marginal Likelihood P(E)**	**Likelihood P(E|s)**	**Bayes Factor BF(s,E)**	**Generalized BF GBF(s,E)**
A priori		73.7569%	100.0000%			
Fib_Onset	>258 (4/4)	83.8710%	51.3812%	58.4270%	1.1371	1.3298
Ferr_Onset	<=12095.848 (1/3)	87.7397%	37.7105%	44.8597%	1.1896	1.3438
Na_Onset	>131.45 (2/2)	90.4539%	31.0990%	38.1391%	1.2264	1.3659
AST_Onset	<=222.406 (1/3)	93.0291%	21.5178%	27.1402%	1.2613	1.3586
BUN_Onset	<=37.66 (1/4)	94.7301%	16.7785%	21.5495%	1.2844	1.3625
Crea_Onset	<=0.684 (1/4)	96.1098%	12.2019%	15.8998%	1.3031	1.3604

Other Nodes		
Node	**Prior Value/Mean**	**Posterior Value/Mean**
Dis_dur_at_MAS_onset	1.9479	1.8033
Neut_Onset	8.2470	7.9732
PLT_Onset	214.4928	226.8775
ESR_Onset	51.1492	54.0548
ALT_Onset	257.1122	114.6053
LDH_Onset	2,281.0383	2,076.9366
Bili_Onset	1.4038	1.2994
Album_Onset	3.1603	3.2147
Ddi_Onset	6,083.3248	5,748.6188
ClusterProb1	0.2448	0.2185

Tab 5. – The combination of values with the highest the probability of absence of a severe outcome

Severe Course = 1						
Node	**Evidence**	**Posterior Probability P(s|E)**	**Marginal Likelihood P(E)**	**Likelihood P(E|s)**	**Bayes Factor BF(s,E)**	**Generalized BF GBF(s,E)**
A priori		26.2431%	100.0000%			
Fib_Onset	<=115.5 (1/4)	59.1837%	13.5359%	30.5263%	2.2552	2.8067
Ferr_Onset	>13382.878 (3/3)	74.6402%	4.4061%	12.5319%	2.8442	3.1084
Na_Onset	<=131.45 (1/2)	87.5487%	1.3444%	4.4851%	3.3361	3.4458
ESR_Onset	<=31 (1/4)	91.7642%	0.7156%	2.5022%	3.4967	3.5608
AST_Onset	>345.269 (3/3)	95.8632%	0.2668%	0.9745%	3.6529	3.6790

Other Nodes		
Node	**Prior Value/Mean**	**Posterior Value/Mean**
Dis_dur_at_MAS_onset	1.9479	2.3983
Neut_Onset	8.2470	9.0999
PLT_Onset	214.4928	175.9195
ALT_Onset	257.1122	610.3555
LDH_Onset	2,281.0383	2,916.7316
Bili_Onset	1.4038	1.7291
Album_Onset	3.1603	2.9909
Ddi_Onset	6,083.3248	7,125.7972
BUN_Onset	30.9457	38.0121

Tab 6. – The combination of values with the highest the probability of occurrence of a severe outcome

Search Method: Hard Evidence

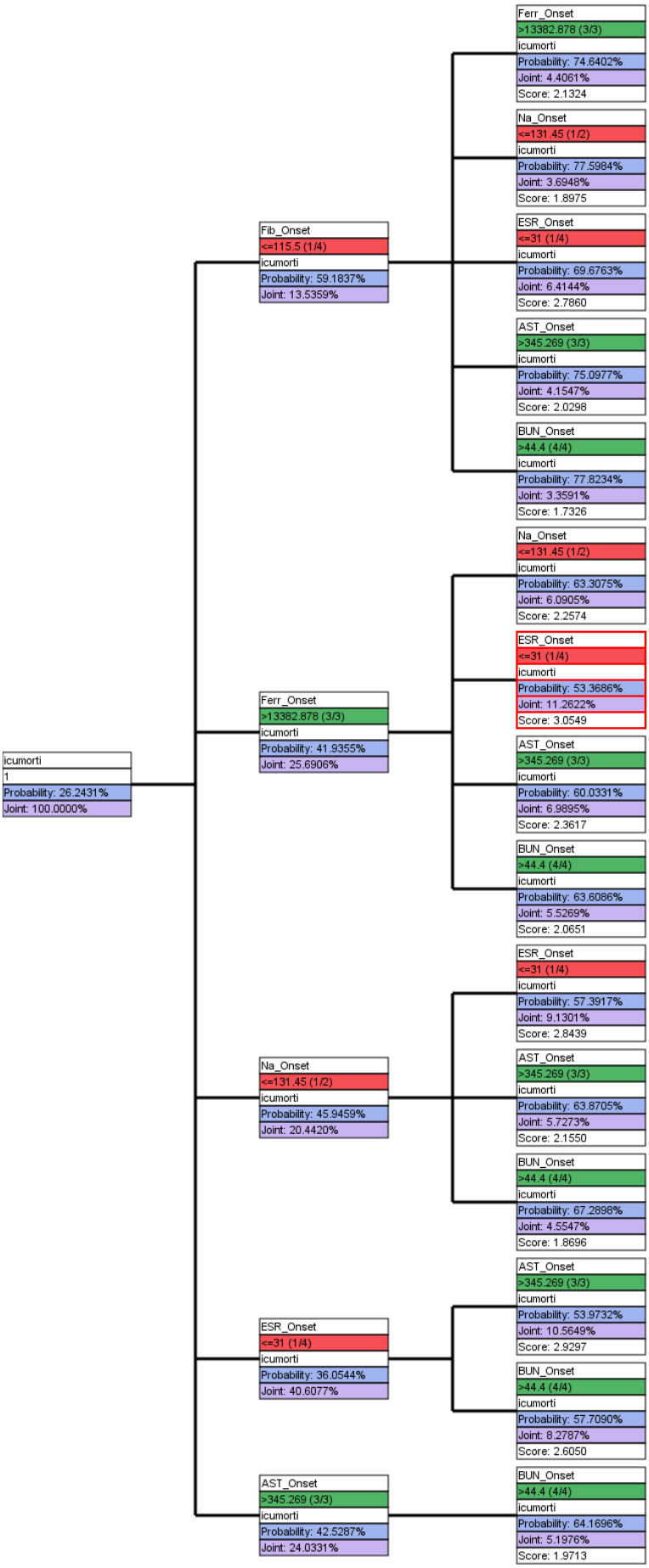


Fig. 7 – target optimization tree for severe outcome. The tree identifies laboratory patterns associated with distinct probabilities of severe course

The Target Optimization Tree in Fig. 7 describes sequence of observations associated with different probability of severe course. The combination of fibrinogen values < 115 mg/dl and Na levels < 131 mg/dl shows the highest probability (77,5%) of a severe outcome. The tree Is based on a numeric score for each variable, derived from a genetic algorithm, which is reported in Fig. 8. The score is obtained by weighting the importance of the node, based on its effect on outcome and joint probabilities, and their position in the tree. In this way the score quantifies the relevance of specific predictors values. Evidences with the highest score are those associated with the highest probability of the outcome.

Node	State	Score
Ferr_Onset	>13382.878 (3/3)	10.8054
Fib_Onset	<=115.5 (1/4)	10.5783
ESR_Onset	<=31 (1/4)	9.8770
Na_Onset	<=131.45 (1/2)	8.9459
AST_Onset	>345.269 (3/3)	6.7094
BUN_Onset	>44.4 (4/4)	5.1218

Fig. 8 – Optimization score for the most predictive values for severe outcome

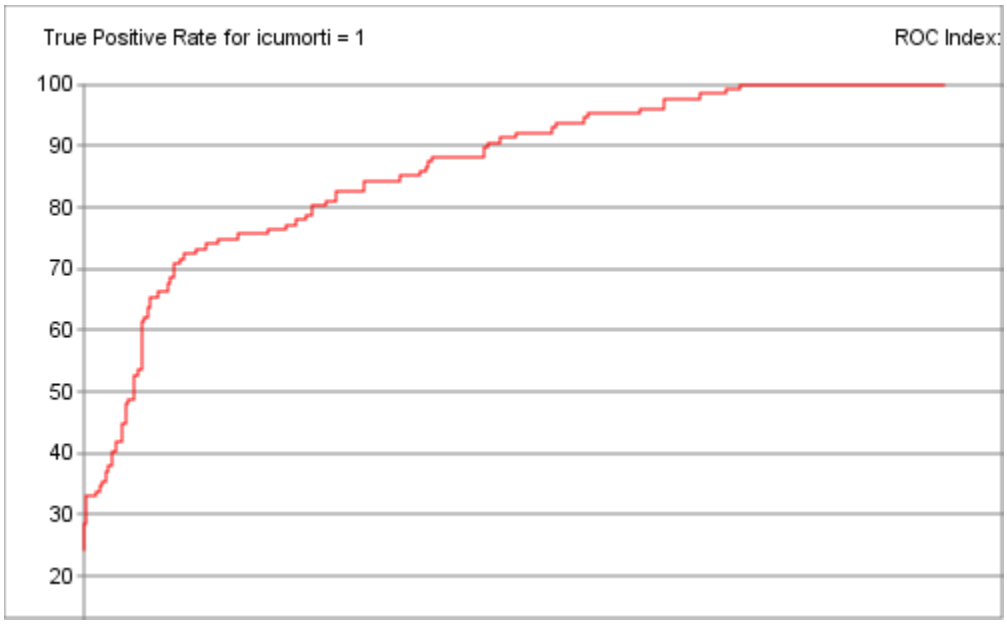
3.1.7 Accuracy assessment

We evaluated the accuracy of the TAN and its performance through the AUC (c-index) of its ROC curves. The model for severe course achieved a c-index of 0.828. Adding the obtained latent

clusters to the BN model increased only slightly the prediction accuracy, up to a c-index of 0.834.

Accuracy metrics and ROC curve of the basic model is reported in Fig. 9 e Tab 7.

Fig. 9 – ROC curve for the occurrence of severe outcome



Tab 7 – Accuracy indexes for the ROC curve for severe outcome.

Target: Severe Course		
Value	0	1
Gini Index	16.6963%	49.0061%
Relative Gini Index	65.6965%	65.7045%
Lift Index	1.2209	1.8690
Relative Lift Index	94.4399%	78.9984%
ROC Index	82.8502%	82.8543%
Calibration Index	74.6629%	75.6270%
Binary Log-Loss	0.4785	0.4785

Overall Precision	82.5967%
Mean Precision	74.3599%
Overall Reliability	81.8722%
Mean Reliability	77.5735%
Overall Relative Gini Index	65.6985%
Mean Relative Gini Index	65.7005%
Overall Relative Lift Index	90.5156%
Mean Relative Lift Index	86.7192%
Overall ROC Index	82.8513%
Mean ROC Index	82.8523%
Overall Calibration Index	74.9079%
Mean Calibration Index	75.1450%
Overall Log-Loss	0.4785
Mean Binary Log-Loss	0.4785

3.2 Mortality model

3.2.1 Structure and Parameters of the Bayesian Networks

The BN graph for mortality prediction is shown in Fig. 10. Ferritin, triglycerides, LDH, sodium, PLT, neutrophil count, creatinine, BUN, albumin and fibrinogen levels resulted predictive of death risk. The probability of fatal outcome depending on different values of the prognostic variable are reported in Fig.11.

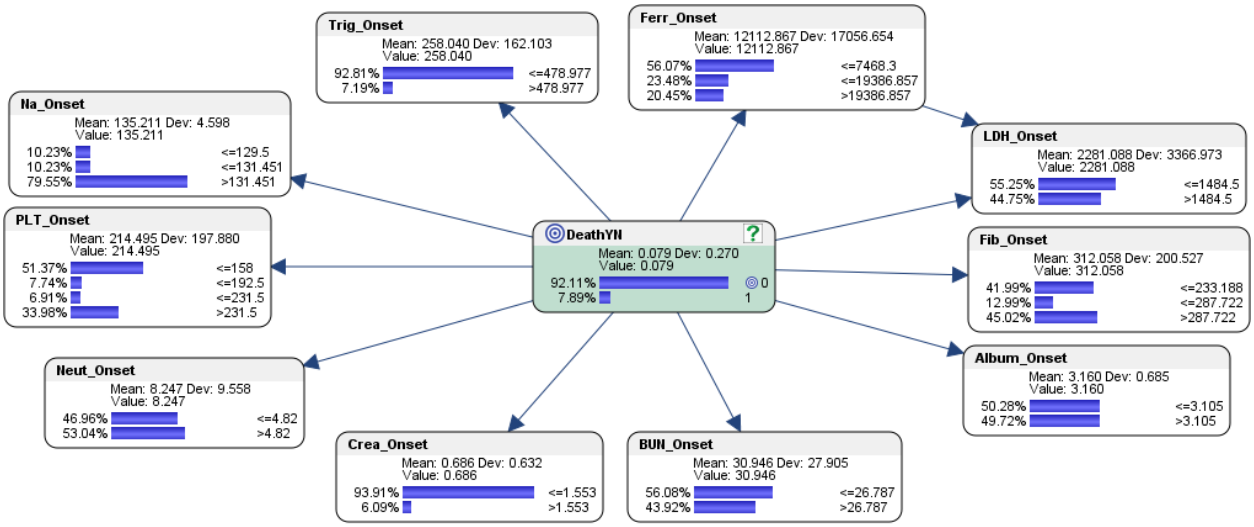


Fig. 10 – The structure of the BN model for mortality prediction

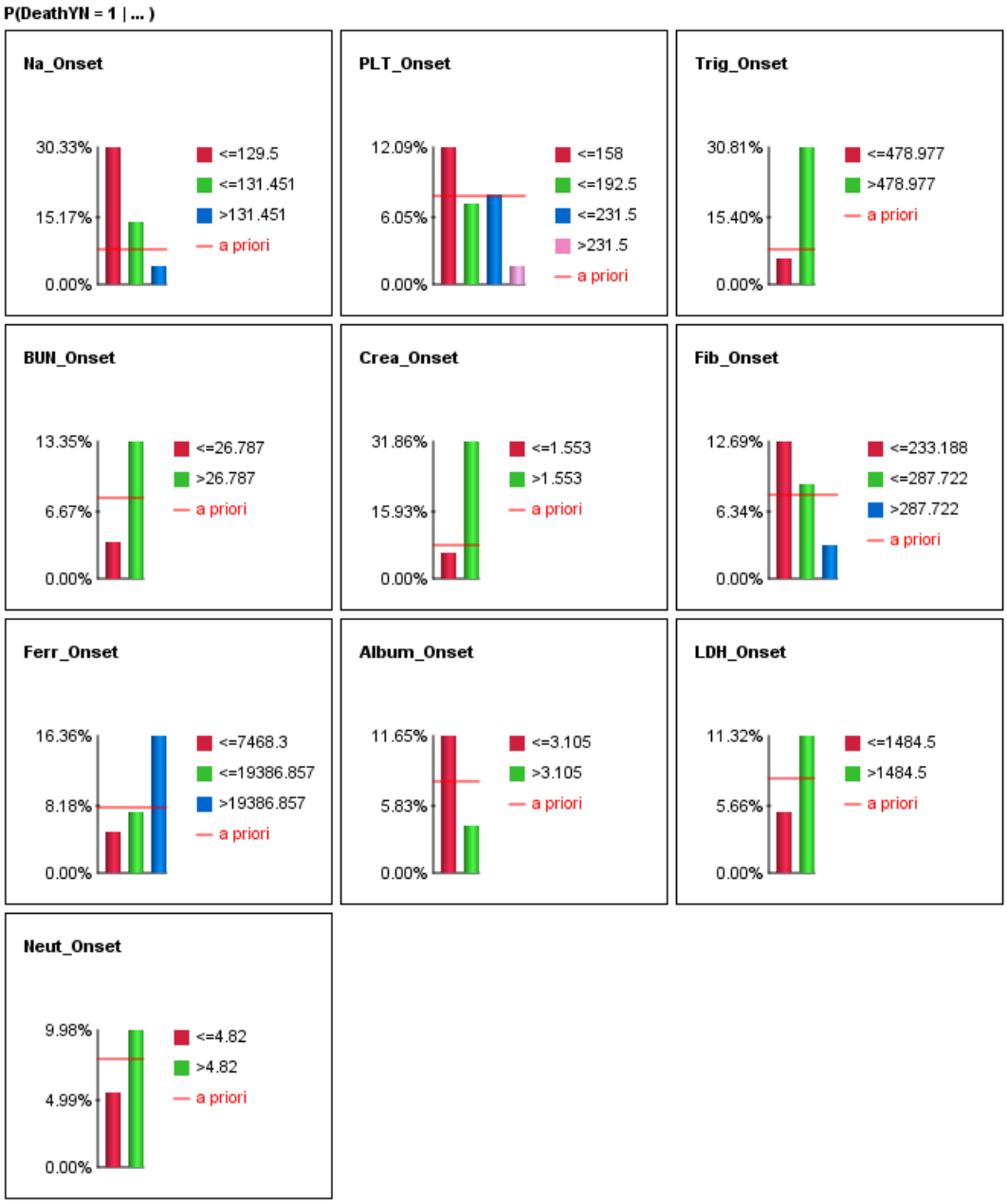


Fig. 11 – Histograms showing the probability of death according to the values of each predictor

3.2.1 Importance ranking of prognostic variables:

Effects of variables on mortality: Which are the causal drivers of fatal outcome?

As reported in Tab. 8 and Fig. 12, serum sodium, triglycerides, creatinine and BUN values had the highest effect of mortality.

Total Effects on Target DeathYN									
Node	Prior Value/Mean	Standardized Total Effects	Total Effects	G-test	df	p-value	G-test (Data)	df (Data)	p-value (Data)
Na_Onset	135.2105	-0.2942	-0.0230	23.8603	2	0.0007%	26.3253	2	0.0002%
Trig_Onset	258.0396	0.2368	0.0006	13.1256	1	0.0291%	12.8198	1	0.0343%
Crea_Onset	0.6863	0.2265	0.1195	11.7619	1	0.0605%	11.4903	1	0.0700%
BUN_Onset	30.9463	0.1794	0.0029	11.8143	1	0.0588%	10.4840	1	0.1204%
PLT_Onset	214.4953	-0.1739	-0.0003	13.4351	3	0.3784%	14.0501	3	0.2838%
Fib_Onset	312.0583	-0.1655	-0.0003	10.6279	2	0.4923%	8.7256	2	1.2742%
Ferr_Onset	12,112.8674	0.1628	0.0000	8.3235	2	1.5580%	10.1497	2	0.6252%
Album_Onset	3.1603	-0.1405	-0.0735	7.4357	1	0.6394%	6.3928	1	1.1459%
LDH_Onset	2,281.0878	0.1147	0.0000	4.7530	1	2.9248%	5.4993	1	1.9024%
Neut_Onset	8.2469	0.0824	0.0039	2.5175	1	11.2589%	3.3043	1	6.9099%

Tab. 8 – Total effects of the different risk factors on mortality

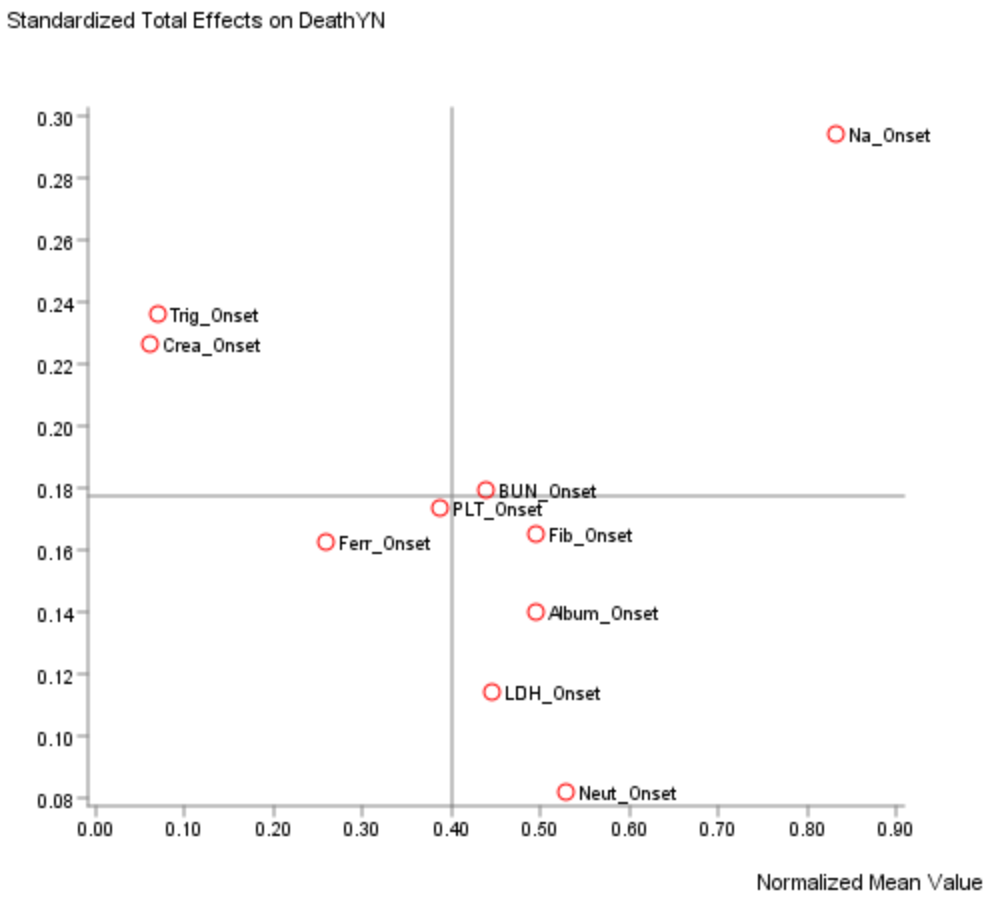
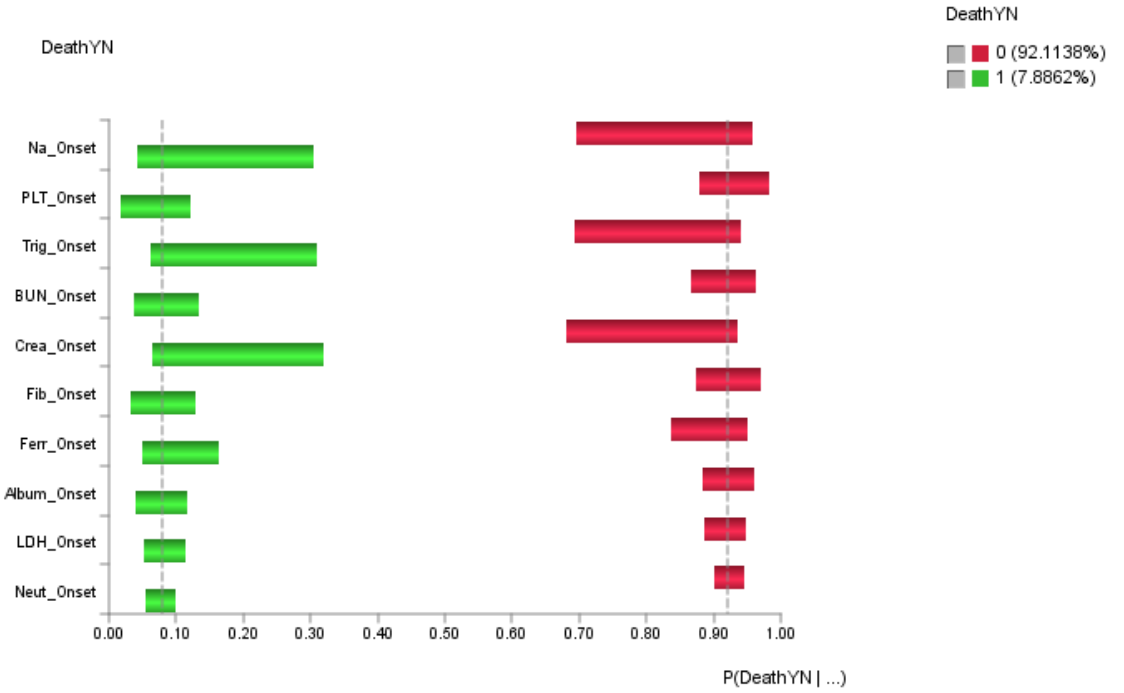
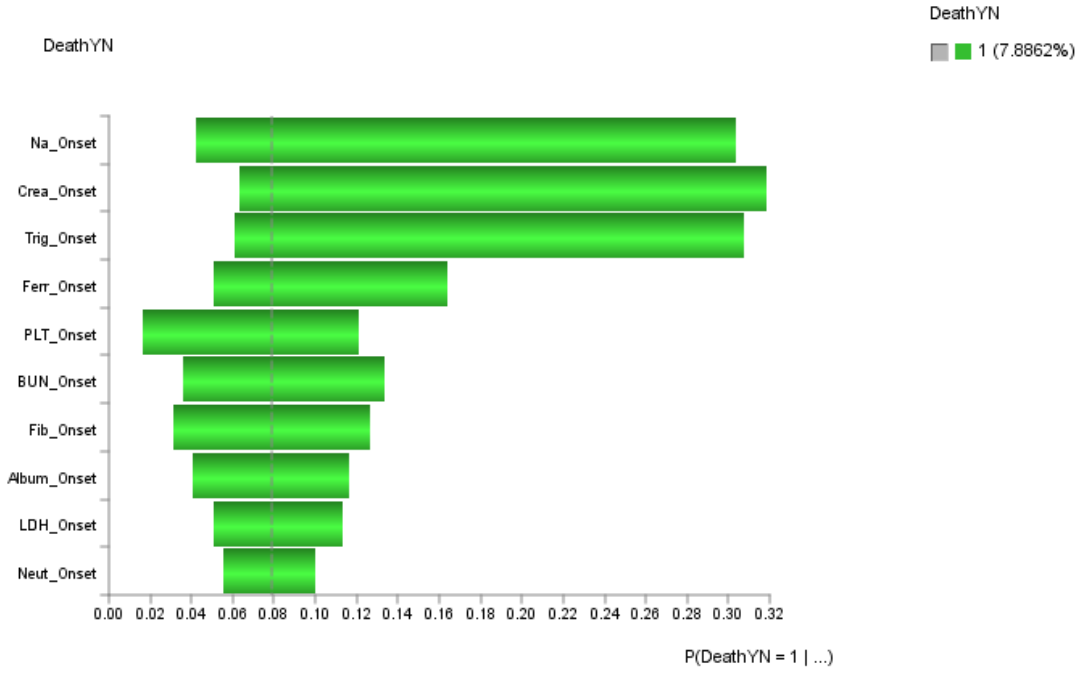


Fig. 12 – Standardized total effect of each variable, normalized for its mean value

3.2.2 Sensitivity analysis

Fig. 13 and 14 show how the probability of death highly depends on difference in Na, creatinine





and triglycerides values.

Fig. 13 and 14 – Tornado plots of both state of the outcome (Fig 13) and for death (Fig 14) from sensitivity analyses in the model for mortality

3.2.3 Information theoretical metrics: Which are the most informative variables for mortality?

As highlighted in Tab. 9 and Fig. 15, serum Na concentration shows the higher mutual information, followed by PLT and triglycerides levels.

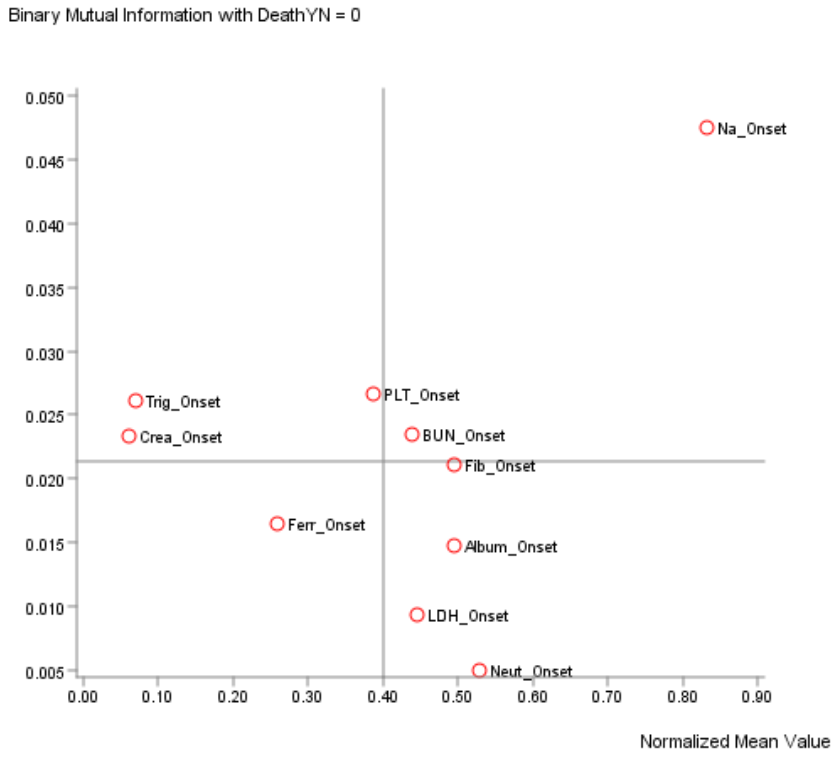
Overall Analysis with DeathYN											
Node	Mutual Information	Normalized Mutual Information	Relative Mutual Information	Relative Significance	Prior Mean Value	G-test	df	p-value	G-test (Data)	df (Data)	p-value (Data)
Na_Onset	0.0475	4.7546%	11.9415%	1.0000	135.2105	23.8603	2	0.0007%	26.3253	2	0.0002%
PLT_Onset	0.0268	2.6772%	6.7239%	0.5631	214.4953	13.4351	3	0.3784%	14.0501	3	0.2838%
Trig_Onset	0.0262	2.6155%	6.5691%	0.5501	258.0396	13.1256	1	0.0291%	12.8198	1	0.0343%
BUN_Onset	0.0235	2.3542%	5.9128%	0.4951	30.9463	11.8143	1	0.0588%	10.4840	1	0.1204%
Crea_Onset	0.0234	2.3438%	5.8865%	0.4929	0.6863	11.7619	1	0.0605%	11.4903	1	0.0700%
Fib_Onset	0.0212	2.1178%	5.3190%	0.4454	312.0583	10.6279	2	0.4923%	8.7256	2	1.2742%
Ferr_Onset	0.0166	1.6586%	4.1657%	0.3488	12,112.8674	8.3235	2	1.5580%	10.1497	2	0.6252%
Album_Onset	0.0148	1.4817%	3.7214%	0.3116	3.1603	7.4357	1	0.6394%	6.3928	1	1.1459%
LDH_Onset	0.0095	0.9471%	2.3787%	0.1992	2,281.0878	4.7530	1	2.9248%	5.4993	1	1.9024%
Neut_Onset	0.0050	0.5017%	1.2599%	0.1055	8.2469	2.5175	1	11.2589%	3.3043	1	6.9099%

Tab. 9 – MI for each variable in the mortality model

Local Analyzes with Target States										
DeathYN = 0 (92.1138%)										
Node	Binary Mutual Information	Relative Binary Mutual Information	Binary Relative Significance	Posterior Mean Value	Max Bayes Factor			Min Bayes Factor		
Na_Onset	0.0475	11.9415%	1.0000	135.5073	>131.451 (3/3)	82.7256%	1.0400	<=129.5 (1/3)	7.7354%	0.7563
PLT_Onset	0.0268	6.7239%	0.5631	222.2981	>231.5 (4/4)	36.2732%	1.0676	<=158 (1/4)	49.0286%	0.9543
Trig_Onset	0.0262	6.5691%	0.5501	250.4107	<=478.977 (1/2)	94.5959%	1.0193	>478.977 (2/2)	5.4041%	0.7512
BUN_Onset	0.0235	5.9128%	0.4951	30.0809	<=26.787 (1/2)	58.6809%	1.0465	>26.787 (2/2)	41.3191%	0.9407
Crea_Onset	0.0234	5.8865%	0.4929	0.6524	<=1.553 (1/2)	95.4953%	1.0169	>1.553 (2/2)	4.5047%	0.7397
Fib_Onset	0.0212	5.3190%	0.4454	320.0748	>287.722 (3/3)	47.3449%	1.0515	<=233.188 (1/3)	39.7978%	0.9479
Ferr_Onset	0.0166	4.1657%	0.3488	11,424.4631	<=7468.3 (1/3)	57.8001%	1.0308	>19386.857 (3/3)	18.5652%	0.9080
Album_Onset	0.0148	3.7214%	0.3116	3.1815	>3.105 (2/2)	51.7788%	1.0413	<=3.105 (1/2)	48.2212%	0.9591
LDH_Onset	0.0095	2.3787%	0.1992	2,224.1180	<=1484.5 (1/2)	56.9155%	1.0302	>1484.5 (2/2)	43.0845%	0.9627
Neut_Onset	0.0050	1.2599%	0.1055	8.1102	<=4.82 (1/2)	48.1657%	1.0256	>4.82 (2/2)	51.8343%	0.9773
DeathYN = 1 (7.8862%)										
Node	Binary Mutual Information	Relative Binary Mutual Information	Binary Relative Significance	Posterior Mean Value	Max Bayes Factor			Min Bayes Factor		
Na_Onset	0.0475	11.9415%	1.0000	131.7441	<=129.5 (1/3)	39.3351%	3.8461	>131.451 (3/3)	42.3971%	0.5330
PLT_Onset	0.0268	6.7239%	0.5631	123.3542	<=158 (1/4)	78.7680%	1.5332	>231.5 (4/4)	7.1368%	0.2101
Trig_Onset	0.0262	6.5691%	0.5501	347.1489	>478.977 (2/2)	28.1029%	3.9064	<=478.977 (1/2)	71.8971%	0.7747
BUN_Onset	0.0235	5.9128%	0.4951	41.0547	>26.787 (2/2)	74.3545%	1.6928	<=26.787 (1/2)	25.6455%	0.4573
Crea_Onset	0.0234	5.8865%	0.4929	1.0816	>1.553 (2/2)	24.6010%	4.0399	<=1.553 (1/2)	75.3990%	0.8029
Fib_Onset	0.0212	5.3190%	0.4454	218.4218	<=233.188 (1/3)	67.5521%	1.6089	>287.722 (3/3)	17.9203%	0.3980

Ferr_Onset	0.0166	4.1657%	0.3488	20,153.7356	>19386.857 (3/3)	42.4094%	2.0743	<=7468.3 (1/3)	35.8747%	0.6398
Album_Onset	0.0148	3.7214%	0.3116	2.9130	<=3.105 (1/2)	74.2794%	1.4774	>3.105 (2/2)	25.7206%	0.5173
LDH_Onset	0.0095	2.3787%	0.1992	2,946.5205	>1484.5 (2/2)	64.2396%	1.4354	<=1484.5 (1/2)	35.7604%	0.6473
Neut_Onset	0.0050	1.2599%	0.1055	9.8446	>4.82 (2/2)	67.0958%	1.2651	<=4.82 (1/2)	32.9042%	0.7007

Tab. 10 – MI for each variable in the mortality model for each outcome state



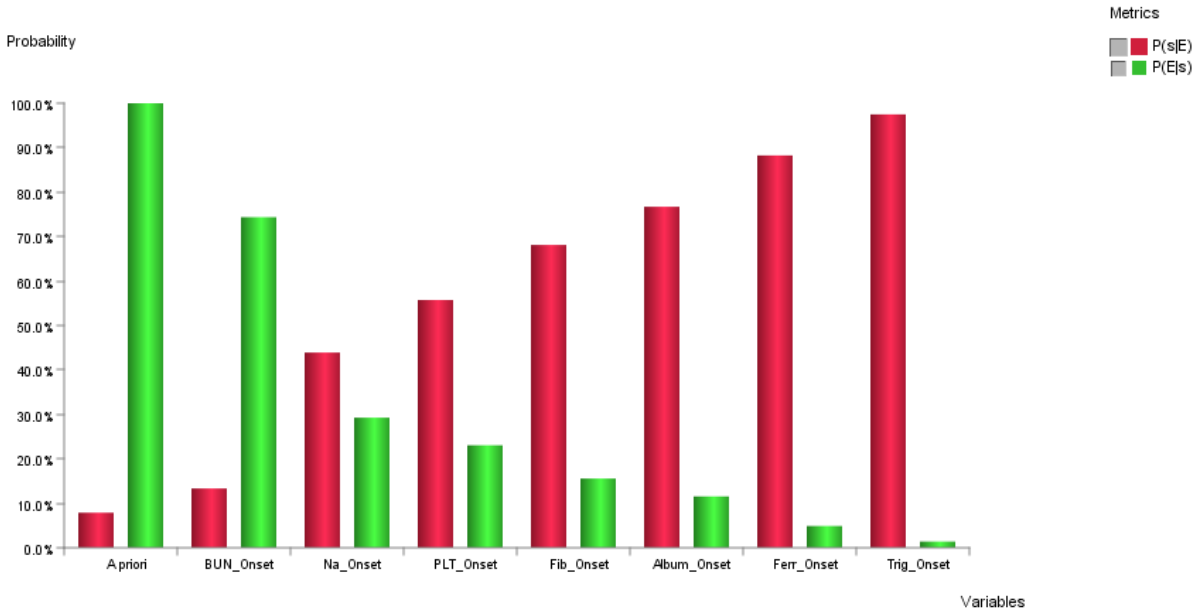


Fig. 15 – MI for each variable, according to mean values, in the mortality model

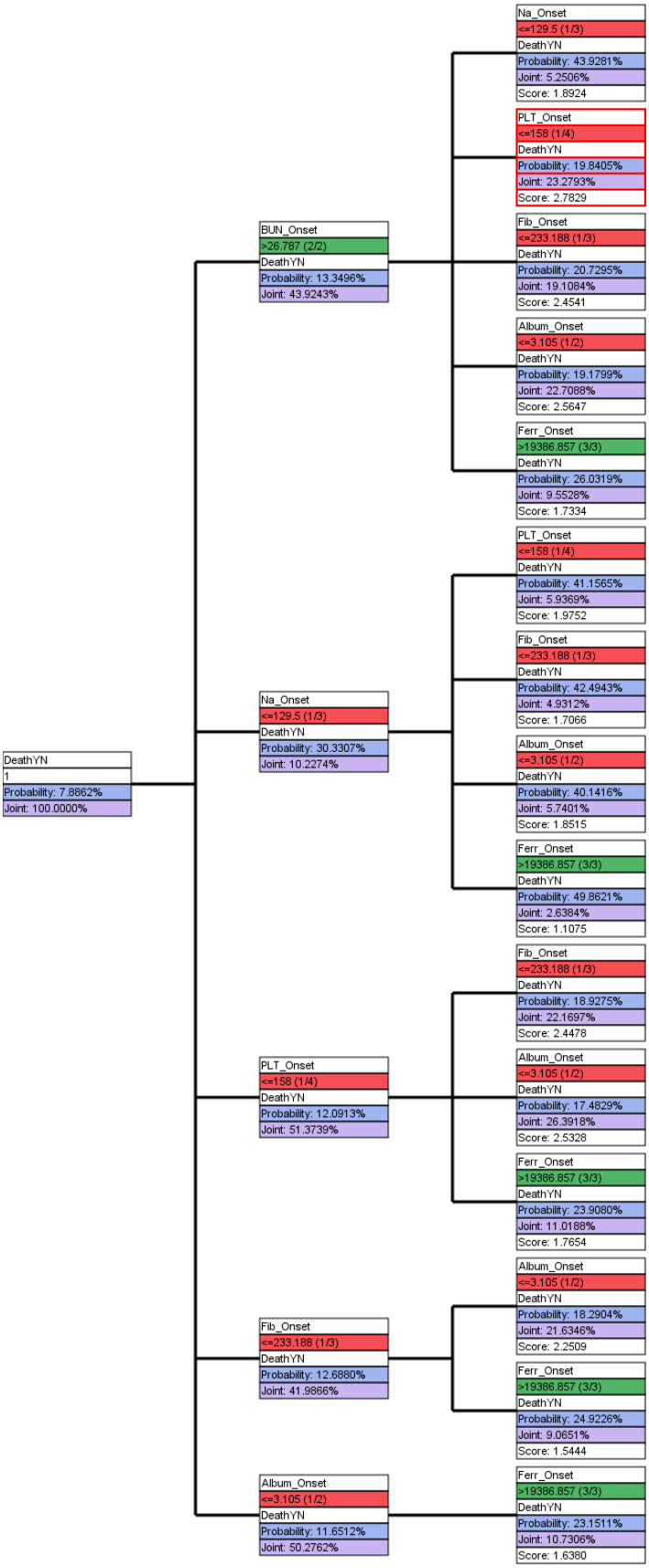
3.2.4 Target Dynamic Profile: Which combination of values are associated with the highest probability of death?

As shown in Tab. 11 and in the target optimization tree for mortality, the laboratory pattern that show the highest probability of a fatal outcome is defined by BUN, Na, PLT, fibrinogen, albumin, ferritin and triglycerides values. Sodium levels above 131 mg/dl are associated with an elevated (95,76%) probability of survival. The target optimization tree reporting different death probabilities according to the combination of specific evidences and the corresponding optimization score are reported in Fig. 16

Tab. 11 – Optimized laboratory profiles that identify the maximum probability of survival and morrtality

Dynamic Profile DeathYN: Probability Maximization (A posteriori)						
Search Method: Hard Evidence						
DeathYN = 0						
Node	**Evidence**	**Posterior Probability P(s|E)**	**Marginal Likelihood P(E)**	**Likelihood P(E|s)**	**Bayes Factor BF(s,E)**	**Generalized BF GBF(s,E)**
A priori		92.1138%	100.0000%			
Na_Onset	>131.451 (3/3)	95.7967%	79.5452%	82.7256%	1.0400	1.2315
Other Nodes						
Node		**Prior Value/Mean**		**Posterior Value/Mean**		
Neut_Onset		8.2469		8.1831		
PLT_Onset		214.4953		218.1393		
LDH_Onset		2,281.0878		2,254.4825		
Ferr_Onset		12,112.8674		11,791.3780		
Trig_Onset		258.0396		254.4769		
Album_Onset		3.1603		3.1702		
Fib_Onset		312.0583		315.8021		
BUN_Onset		30.9463		30.5421		
Crea_Onset		0.6863		0.6705		
DeathYN = 1						
Node	**Evidence**	**Posterior Probability P(s|E)**	**Marginal Likelihood P(E)**	**Likelihood P(E|s)**	**Bayes Factor BF(s,E)**	**Generalized BF GBF(s,E)**
A priori		7.8862%	100.0000%			
BUN_Onset	>26.787 (2/2)	13.3496%	43.9243%	74.3545%	1.6928	3.7014
Na_Onset	<=129.5 (1/3)	43.9281%	5.2506%	29.2474%	5.5703	7.4595
PLT_Onset	<=158 (1/4)	55.7253%	3.2602%	23.0376%	7.0662	8.8820
Fib_Onset	<=233.188 (1/3)	68.1160%	1.8017%	15.5624%	8.6374	10.0450
Album_Onset	<=3.105 (1/2)	76.6945%	1.1886%	11.5596%	9.7252	10.8656
Ferr_Onset	>19386.857 (3/3)	88.2594%	0.4380%	4.9024%	11.1917	11.7171
Trig_Onset	>478.977 (2/2)	97.5058%	0.1114%	1.3777%	12.3642	12.5229
Other Nodes						
Node		**Prior Value/Mean**		**Posterior Value/Mean**		
Neut_Onset		8.2469		9.8013		
LDH_Onset		2,281.0878		3,318.5355		
Crea_Onset		0.6863		1.0709		

Search Method: Hard Evidence

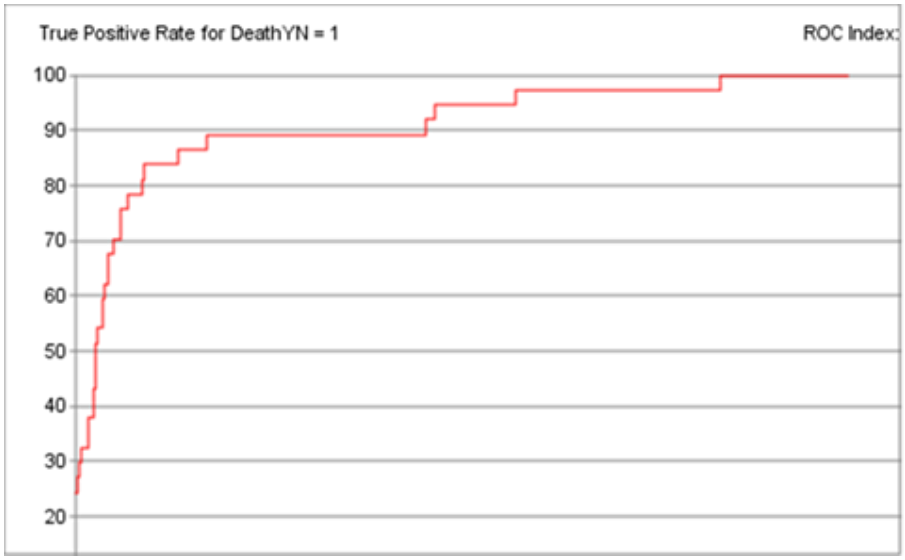


Node	State	Score
BUN_Onset	>26.787 (2/2)	26.2646
PLT_Onset	<=158 (1/4)	13.9426
Na_Onset	<=129.5 (1/3)	13.6540
Fib_Onset	<=233.188 (1/3)	9.8038
Album_Onset	<=3.105 (1/2)	8.0395
LDH_Onset	>1484.5 (2/2)	5.8462
Ferr_Onset	>19386.857 (3/3)	5.6958
Neut_Onset	>4.82 (2/2)	5.0803
Trig_Onset	>478.977 (2/2)	3.7946
Crea_Onset	>1.553 (2/2)	2.9961

Fig. 16 – Target Optimization Tree and optimization score for the mortality model

3.2.4 Accuracy assessment

The model for mortality showed a mean ROC index of 0.816. A difference in the ability of the TAN to forecast the occurrence of death ore survival was observed, with a higher ROC index for death (87.6%) than for survival (75.6%)- ROC curves and accuracy metrics are reported in Fig. 17 and Tab. 12.



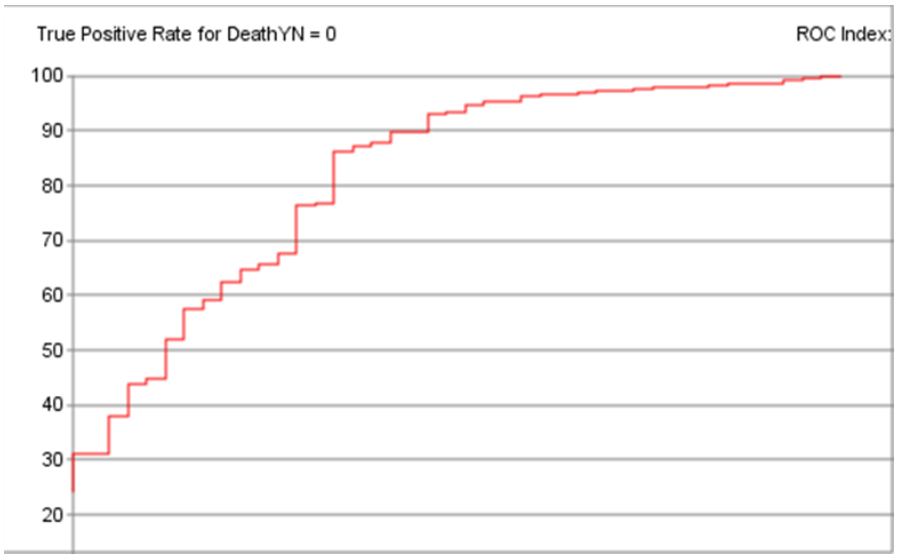


Fig. 17 – ROC curves for mortality prediction

Overall Precision	92.2636%
Mean Precision	71.3451%
Overall Reliability	88.5070%
Mean Reliability	71.8457%
Overall Relative Gini Index	53.2842%
Mean Relative Gini Index	63.3316%
Overall Relative Lift Index	93.5599%
Mean Relative Lift Index	83.6636%
Overall ROC Index	76.6456%
Mean ROC Index	81.6677%
Overall Calibration Index	73.1542%
Mean Calibration Index	71.0768%
Overall Log-Loss	0.2157
Mean Binary Log-Loss	0.2993

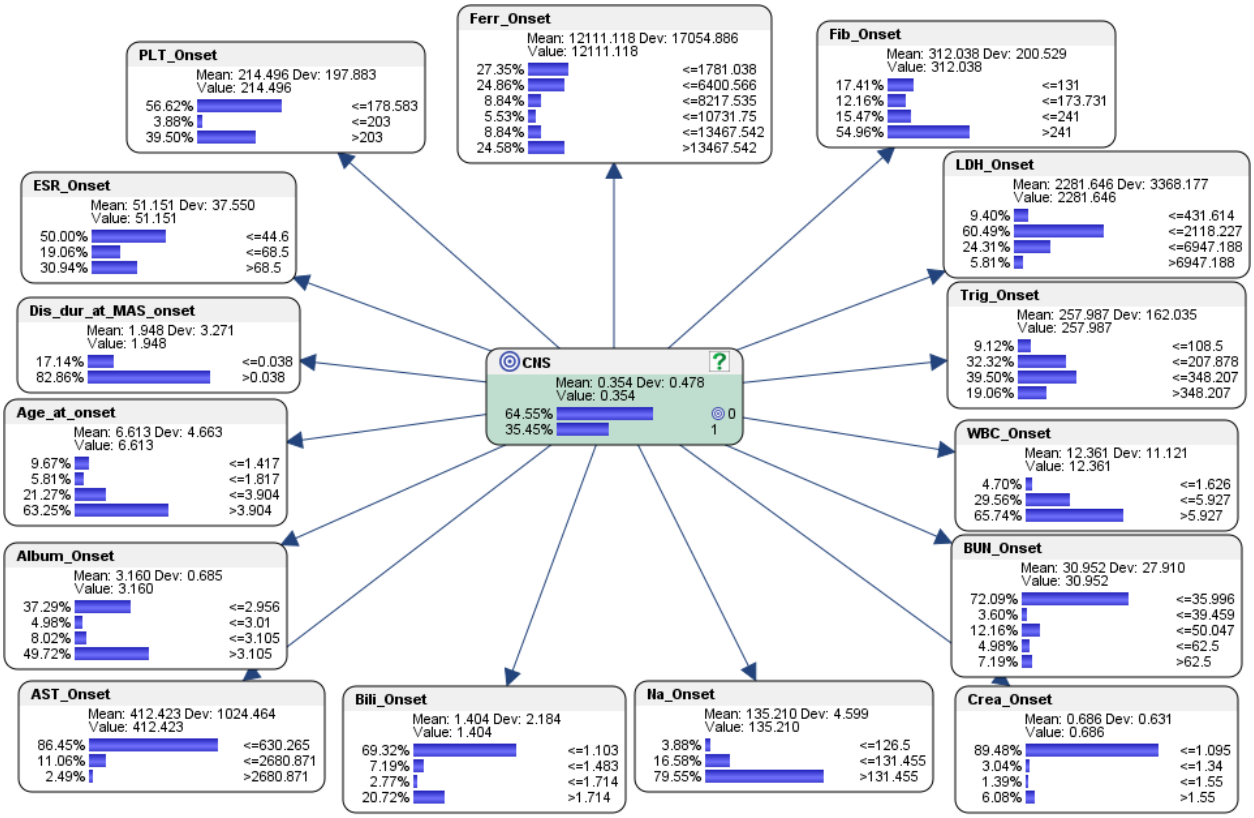
Target: DeathYN		
Value	0	1
Gini Index	5.8175%	69.4751%
Relative Gini Index	51.3639%	75.2994%
Lift Index	1.0691	2.5466
Relative Lift Index	95.4513%	71.8758%
ROC Index	75.6857%	87.6497%
Calibration Index	73.5512%	68.6024%
Binary Log-Loss	0.3814	0.2171

Tab. 12 – Accuracy of the mortality model

3.3 CNS disease model

3.3.1 Structure of the Bayesian Network

As represented in Fig. 18, the model for CNS involvement retained 15 predictors, including age at SJIA onset and disease duration at MAS diagnosis.


PLT_Onset
Mean: 214.496 Dev: 197.883
Value: 214.496
56.62% <=178.583
3.88% <=203
39.50% >203
Ferr_Onset
Mean: 12111.118 Dev: 17054.886
Value: 12111.118
27.35% <=1781.038
24.86% <=6400.566
8.84% <=8217.535
5.53% <=10731.75
8.84% <=13467.542
24.58% >13467.542
Fib_Onset
Mean: 312.038 Dev: 200.529
Value: 312.038
17.41% <=131
12.16% <=173.731
15.47% <=241
54.96% >241
ESR_Onset
Mean: 51.151 Dev: 37.550
Value: 51.151
50.00% <=44.6
19.06% <=68.5
30.94% >68.5
LDH_Onset
Mean: 2281.646 Dev: 3368.177
Value: 2281.646
9.40% <=431.614
60.49% <=2118.227
24.31% <=6947.188
5.81% >6947.188
Dis_dur_at_MAS_onset
Mean: 1.948 Dev: 3.271
Value: 1.948
17.14% <=0.038
82.86% >0.038
CNS
Mean: 0.354 Dev: 0.478
Value: 0.354
64.55% 0
35.45% 1
Trig_Onset
Mean: 257.987 Dev: 162.035
Value: 257.987
9.12% <=108.5
32.32% <=207.878
39.50% <=348.207
19.06% >348.207
Age_at_onset
Mean: 6.613 Dev: 4.663
Value: 6.613
9.67% <=1.417
5.81% <=1.817
21.27% <=3.904
63.25% >3.904
WBC_Onset
Mean: 12.361 Dev: 11.121
Value: 12.361
4.70% <=1.626
29.56% <=5.927
65.74% >5.927
Album_Onset
Mean: 3.160 Dev: 0.685
Value: 3.160
37.29% <=2.956
4.98% <=3.01
8.02% <=3.105
49.72% >3.105
BUN_Onset
Mean: 30.952 Dev: 27.910
Value: 30.952
72.09% <=35.996
3.60% <=39.459
12.16% <=50.047
4.98% <=62.5
7.19% >62.5
AST_Onset
Mean: 412.423 Dev: 1024.464
Value: 412.423
86.45% <=630.265
11.06% <=2680.871
2.49% >2680.871
Bili_Onset
Mean: 1.404 Dev: 2.184
Value: 1.404
69.32% <=1.103
7.19% <=1.483
2.77% <=1.714
20.72% >1.714
Na_Onset
Mean: 135.210 Dev: 4.599
Value: 135.210
3.88% <=126.5
16.58% <=131.455
79.55% >131.455
Crea_Onset
Mean: 0.686 Dev: 0.631
Value: 0.686
89.48% <=1.095
3.04% <=1.34
1.39% <=1.55
6.08% >1.55

P(CNS = 1 | ...)

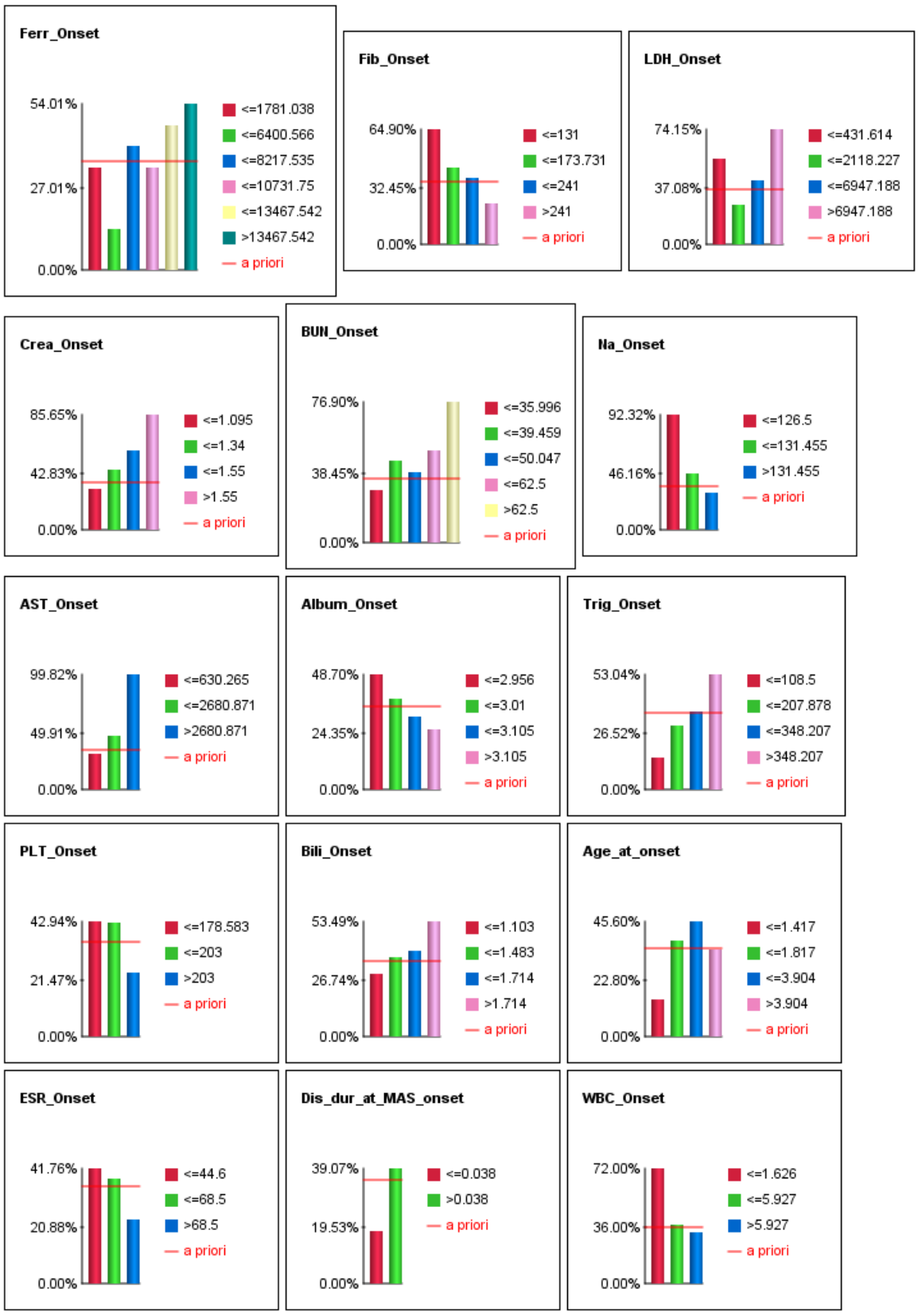

Ferr_Onset
54.01%
27.01%
0.00%
<=1781.038
<=6400.566
<=8217.535
<=10731.75
<=13467.542
>13467.542
a priori
Fib_Onset
64.90%
32.45%
0.00%
<=131
<=173.731
<=241
>241
a priori
LDH_Onset
74.15%
37.08%
0.00%
<=431.614
<=2118.227
<=6947.188
>6947.188
a priori
Crea_Onset
85.65%
42.83%
0.00%
<=1.095
<=1.34
<=1.55
>1.55
a priori
BUN_Onset
76.90%
38.45%
0.00%
<=35.996
<=39.459
<=50.047
<=62.5
>62.5
a priori
Na_Onset
92.32%
46.16%
0.00%
<=126.5
<=131.455
>131.455
a priori
AST_Onset
99.82%
49.91%
0.00%
<=630.265
<=2680.871
>2680.871
a priori
Album_Onset
48.70%
24.35%
0.00%
<=2.956
<=3.01
<=3.105
>3.105
a priori
Trig_Onset
53.04%
26.52%
0.00%
<=108.5
<=207.878
<=348.207
>348.207
a priori
PLT_Onset
42.94%
21.47%
0.00%
<=178.583
<=203
>203
a priori
Bili_Onset
53.49%
26.74%
0.00%
<=1.103
<=1.483
<=1.714
>1.714
a priori
Age_at_onset
45.60%
22.80%
0.00%
<=1.417
<=1.817
<=3.904
>3.904
a priori
ESR_Onset
41.76%
20.88%
0.00%
<=44.6
<=68.5
>68.5
a priori
Dis_dur_at_MAS_onset
39.07%
19.53%
0.00%
<=0.038
>0.038
a priori
WBC_Onset
72.00%
36.00%
0.00%
<=1.626
<=5.927
>5.927
a priori

Fig. 18 – The structure and value distributions of the BN for CNS disease

3.3.2 Importance ranking of prognostic variables:

Effects of variables on mortality: Which are the causal drivers of CNS dysfunction?

Total effects of variables on the probability of CNS disease are reported in Tab.13 and Fig.19.

Fibrinogen, creatinine and BUN exhibit the greatest impact.

Total Effects on Target CNS									
Node	Prior Value/Mean	Standardized Total Effects	Total Effects	G-test	df	p-value	G-test (Data)	df (Data)	p-value (Data)
Fib_Onset	312.0382	-0.3029	-0.0009	36.8757	3	0.0000%	37.2397	3	0.0000%
Crea_Onset	0.6862	0.2798	0.2507	27.9513	3	0.0004%	28.4825	3	0.0003%
BUN_Onset	30.9515	0.2680	0.0056	25.6995	4	0.0036%	25.2046	4	0.0046%
Ferr_Onset	12,111.1177	0.2394	0.0000	37.0235	5	0.0001%	38.5056	5	0.0000%
Na_Onset	135.2102	-0.2386	-0.0328	25.6146	2	0.0003%	26.8922	2	0.0001%
AST_Onset	412.4229	0.2346	0.0001	22.4299	2	0.0013%	22.5069	2	0.0013%
Album_Onset	3.1603	-0.2196	-0.1935	17.8715	3	0.0468%	17.7533	3	0.0494%
LDH_Onset	2,281.6461	0.2123	0.0000	29.5766	3	0.0002%	30.9383	3	0.0001%
Trig_Onset	257.9873	0.2082	0.0007	17.2893	3	0.0616%	17.9314	3	0.0454%
Bili_Onset	1.4040	0.1986	0.0595	13.9618	3	0.2958%	13.6459	3	0.3429%
PLT_Onset	214.4956	-0.1923	-0.0006	13.9902	2	0.0916%	14.9224	2	0.0575%
Dis_dur_at_MAS_onset	1.9477	0.1663	0.0902	10.9543	1	0.0934%	11.4407	1	0.0719%
ESR_Onset	51.1512	-0.1647	-0.0023	10.9712	2	0.4146%	11.5399	2	0.3120%
WBC_Onset	12.3613	-0.1054	-0.0078	10.6112	2	0.4964%	9.8490	2	0.7266%
Age_at_onset	6.6132	0.0080	0.0011	11.0214	3	1.1611%	11.2400	3	1.0496%

Tab. 13 – Total effects on CNS disease

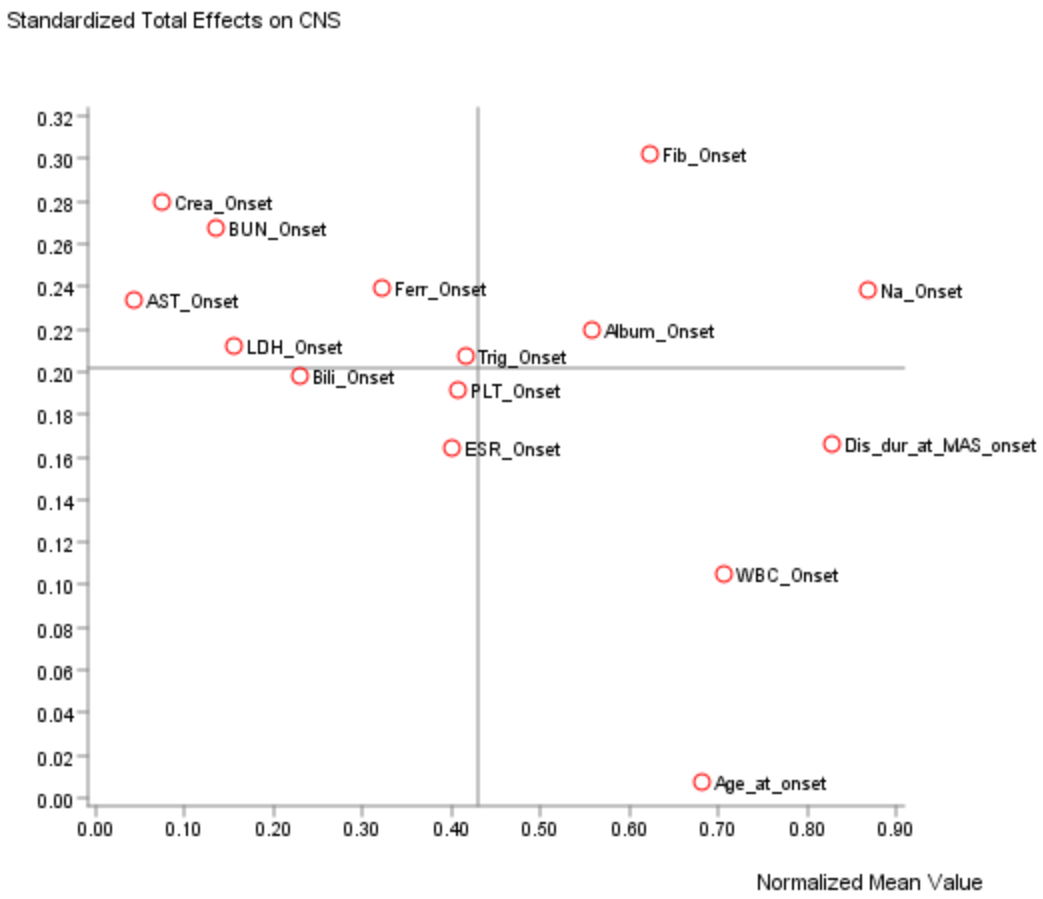
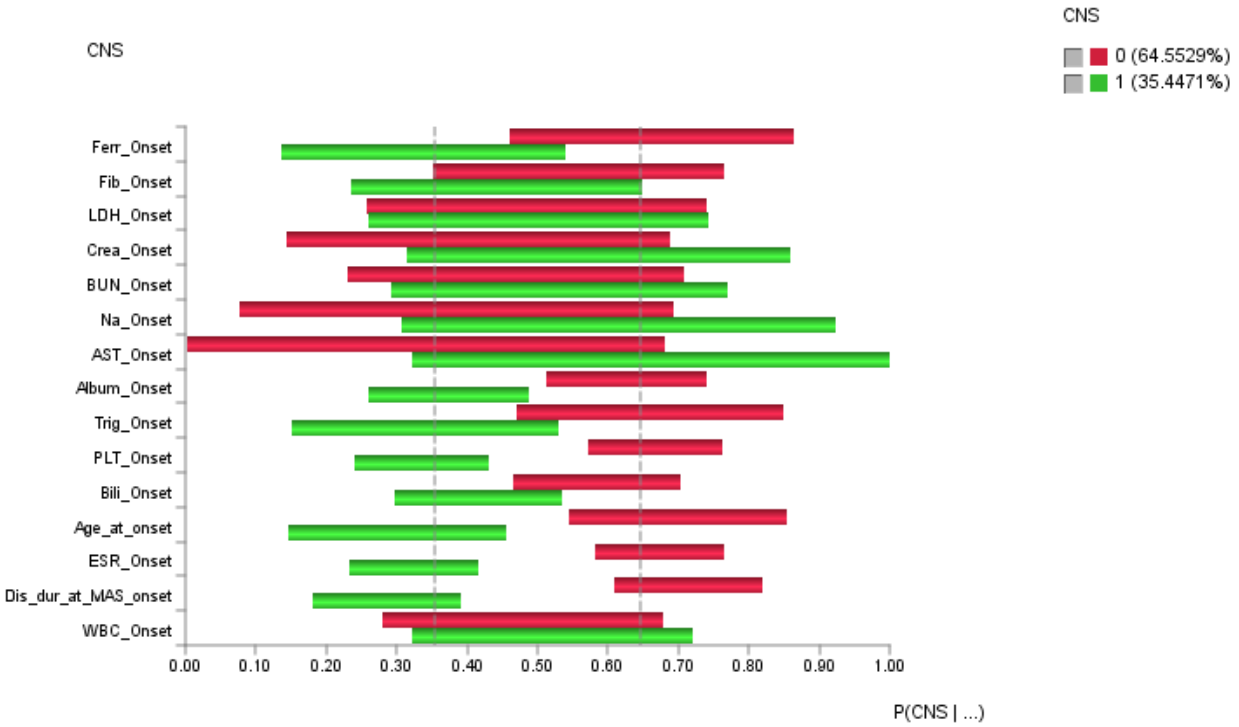


Fig. 19 – Standardized total effects of each predictor on CNS disease

3.3.3 Sensivity analysis

As depicted in the tornado graphs (Fig. 20) AST, Na and creatinine values emerged as the most sensitive parameters in the model for the prediction of CSN disease.


CNS
CNS
0 (64.5529%)
1 (35.4471%)
Ferr_Onset
Fib_Onset
LDH_Onset
Crea_Onset
BUN_Onset
Na_Onset
AST_Onset
Album_Onset
Trig_Onset
PLT_Onset
Bili_Onset
Age_at_onset
ESR_Onset
Dis_dur_at_MAS_onset
WBC_Onset
0.00 0.10 0.20 0.30 0.40 0.50 0.60 0.70 0.80 0.90 1.00
P(CNS | ...)

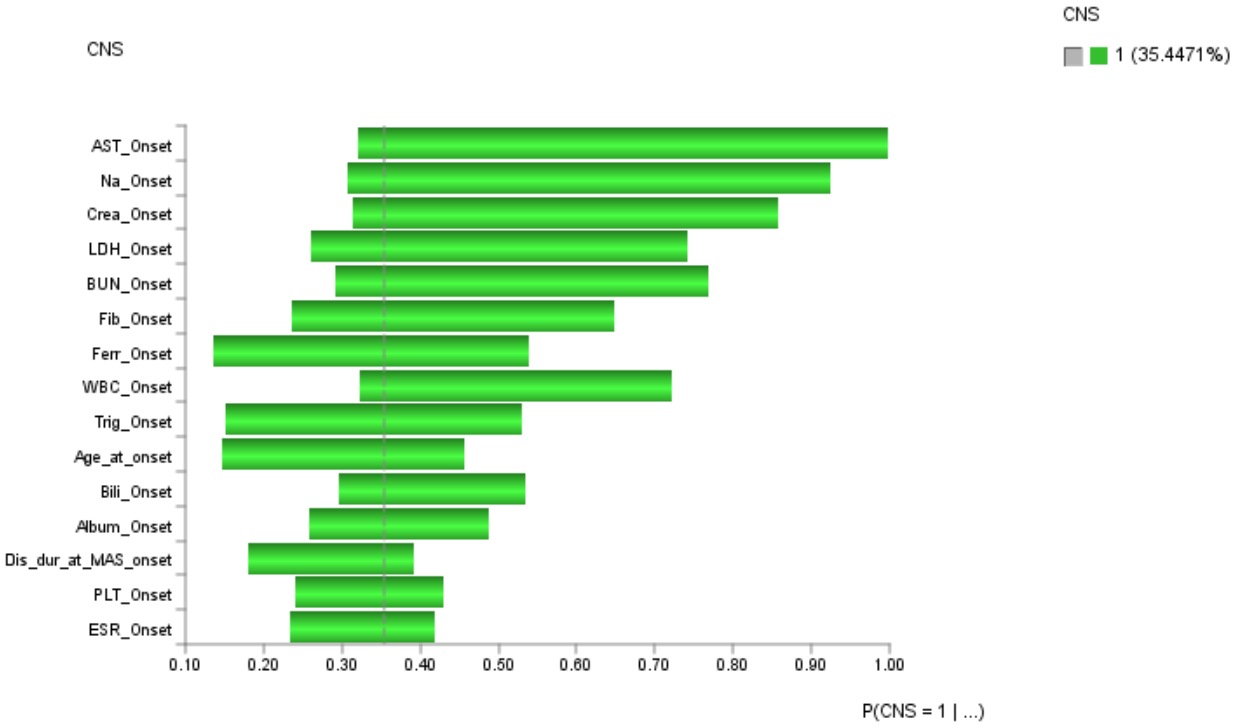

CNS
CNS
1 (35.4471%)
AST_Onset
Na_Onset
Crea_Onset
LDH_Onset
BUN_Onset
Fib_Onset
Ferr_Onset
WBC_Onset
Trig_Onset
Age_at_onset
Bili_Onset
Album_Onset
Dis_dur_at_MAS_onset
PLT_Onset
ESR_Onset
0.10 0.20 0.30 0.40 0.50 0.60 0.70 0.80 0.90 1.00
P(CNS = 1 | ...)

Fig. 20 – Tornado graphs resulting from sensitivity analysis of the CNS model

3.3.4 Information theoretical metrics: Which are the most informative variables for the prediction of CNS disease?

Ferritin, fibrinogen, LDH and creatinine levels are the variables providing the higher information gain in forecasting the occurrence on CNS dysfunction (Tab. 14 and Fig. 21)

Overall Analysis with CNS											
Node	Mutual Information	Normalized Mutual Information	Relative Mutual Information	Relative Significance	Prior Mean Value	G-test	df	p-value	G-test (Data)	df (Data)	p-value (Data)
Ferr_Onset	0.0738	7.3776%	7.8652%	1.0000	12,111.1177	37.0235	5	0.0001%	38.5056	5	0.0000%
Fib_Onset	0.0735	7.3481%	7.8338%	0.9960	312.0382	36.8757	3	0.0000%	37.2397	3	0.0000%
LDH_Onset	0.0589	5.8936%	6.2832%	0.7989	2,281.6461	29.5766	3	0.0002%	30.9383	3	0.0001%
Crea_Onset	0.0557	5.5698%	5.9379%	0.7550	0.6862	27.9513	3	0.0004%	28.4825	3	0.0003%
BUN_Onset	0.0512	5.1211%	5.4596%	0.6941	30.9515	25.6995	4	0.0036%	25.2046	4	0.0046%
Na_Onset	0.0510	5.1042%	5.4415%	0.6918	135.2102	25.6146	2	0.0003%	26.8922	2	0.0001%
AST_Onset	0.0447	4.4695%	4.7650%	0.6058	412.4229	22.4299	2	0.0013%	22.5069	2	0.0013%
Album_Onset	0.0356	3.5612%	3.7966%	0.4827	3.1603	17.8715	3	0.0468%	17.7533	3	0.0494%
Trig_Onset	0.0345	3.4452%	3.6729%	0.4670	257.9873	17.2893	3	0.0616%	17.9314	3	0.0454%
PLT_Onset	0.0279	2.7878%	2.9721%	0.3779	214.4956	13.9902	2	0.0916%	14.9224	2	0.0575%
Bili_Onset	0.0278	2.7821%	2.9660%	0.3771	1.4040	13.9618	3	0.2958%	13.6459	3	0.3429%
Age_at_onset	0.0220	2.1962%	2.3414%	0.2977	6.6132	11.0214	3	1.1611%	11.2400	3	1.0496%
ESR_Onset	0.0219	2.1862%	2.3307%	0.2963	51.1512	10.9712	2	0.4146%	11.5399	2	0.3120%
Dis_dur_at_MAS_onset	0.0218	2.1828%	2.3271%	0.2959	1.9477	10.9543	1	0.0934%	11.4407	1	0.0719%
WBC_Onset	0.0211	2.1145%	2.2542%	0.2866	12.3613	10.6112	2	0.4964%	9.8490	2	0.7266%

Local Analyzes with Target States										
CNS = 0 (64.5529%)										
Node	Binary Mutual Information	Relative Binary Mutual Information	Binary Relative Significance	Posterior Mean Value	Max Bayes Factor			Min Bayes Factor		
Ferr_Onset	0.0738	7.8652%	1.0000	9,637.9148	<=6400.566 (2/6)	33.2654%	1.3381	>13467.542 (6/6)	17.5126%	0.7124
Fib_Onset	0.0735	7.8338%	0.9960	347.0857	>241 (4/4)	65.0345%	1.1832	<=131 (1/4)	9.4634%	0.5437
LDH_Onset	0.0589	6.2832%	0.7989	1,827.2188	<=2118.227 (2/4)	69.3819%	1.1470	>6947.188 (4/4)	2.3248%	0.4004
Crea_Onset	0.0557	5.9379%	0.7550	0.5756	<=1.095 (1/4)	95.2147%	1.0640	>1.55 (4/4)	1.3521%	0.2223
BUN_Onset	0.0512	5.4596%	0.6941	26.3709	<=35.996 (1/5)	79.0644%	1.0968	>62.5 (5/5)	2.5712%	0.3578
Na_Onset	0.0510	5.4415%	0.6918	135.8313	>131.455 (3/3)	85.4738%	1.0745	<=126.5 (1/3)	0.4614%	0.1190
AST_Onset	0.0447	4.7650%	0.6058	246.4138	<=630.265 (1/3)	90.9724%	1.0523	>2680.871 (3/3)	0.0071%	0.0029
Album_Onset	0.0356	3.7966%	0.4827	3.2487	>3.105 (4/4)	57.0949%	1.1484	<=2.956 (1/4)	29.6321%	0.7947
Trig_Onset	0.0345	3.6729%	0.4670	237.3246	<=108.5 (1/4)	11.9842%	1.3140	>348.207 (4/4)	13.8661%	0.7274
PLT_Onset	0.0279	2.9721%	0.3779	235.2325	>203 (3/3)	46.5233%	1.1778	<=178.583 (1/3)	50.0471%	0.8839
Bili_Onset	0.0278	2.9660%	0.3771	1.1689	<=1.103 (1/4)	75.4818%	1.0888	>1.714 (4/4)	14.9293%	0.7205
Age_at_onset	0.0220	2.3414%	0.2977	6.5911	<=1.417 (1/4)	12.7803%	1.3213	<=3.904 (3/4)	17.9253%	0.8427
ESR_Onset	0.0219	2.3307%	0.2963	55.3973	>68.5 (3/3)	36.7241%	1.1869	<=44.6 (1/3)	45.1079%	0.9022
Dis_dur_at_MAS_onset	0.0218	2.3271%	0.2959	1.8390	<=0.038 (1/2)	21.7810%	1.2711	>0.038 (2/2)	78.2190%	0.9439
WBC_Onset	0.0211	2.2542%	0.2866	12.8704	>5.927 (3/3)	69.0367%	1.0502	<=1.626 (1/3)	2.0401%	0.4337

CNS = 1 (35.4471%)										
Node	Binary Mutual Information	Relative Binary Mutual Information	Binary Relative Significance	Posterior Mean Value	Max Bayes Factor			Min Bayes Factor		
Ferr_Onset	0.0738	7.8652%	1.0000	16,615.0805	>13467.542 (6/6)	37.4603%	1.5238	<=6400.566 (2/6)	9.5517%	0.3842
Fib_Onset	0.0735	7.8338%	0.9960	248.2129	<=131 (1/4)	31.8686%	1.8310	>241 (4/4)	36.6248%	0.6663
LDH_Onset	0.0589	6.2832%	0.7989	3,109.2061	>6947.188 (4/4)	12.1467%	2.0919	<=2118.227 (2/4)	44.2897%	0.7322
Crea_Onset	0.0557	5.9379%	0.7550	0.8878	>1.55 (4/4)	14.6973%	2.4163	<=1.095 (1/4)	79.0504%	0.8834
BUN_Onset	0.0512	5.4596%	0.6941	39.2934	>62.5 (5/5)	15.5897%	2.1695	<=35.996 (1/5)	59.3750%	0.8237
Na_Onset	0.0510	5.4415%	0.6918	134.0791	<=126.5 (1/3)	10.0931%	2.6043	>131.455 (3/3)	68.7487%	0.8643
AST_Onset	0.0447	4.7650%	0.6058	714.7431	>2680.871 (3/3)	7.0248%	2.8159	<=630.265 (1/3)	78.2125%	0.9047
Album_Onset	0.0356	3.7966%	0.4827	2.9992	<=2.956 (1/4)	51.2341%	1.3740	>3.105 (4/4)	36.2808%	0.7297
Trig_Onset	0.0345	3.6729%	0.4670	295.6163	>348.207 (4/4)	28.5254%	1.4964	<=108.5 (1/4)	3.9052%	0.4282
PLT_Onset	0.0279	2.9721%	0.3779	176.7316	<=178.583 (1/3)	68.5996%	1.2115	>203 (3/3)	26.7128%	0.6763
Bili_Onset	0.0278	2.9660%	0.3771	1.8321	>1.714 (4/4)	31.2638%	1.5089	<=1.103 (1/4)	58.1121%	0.8383
Age_at_onset	0.0220	2.3414%	0.2977	6.6533	<=3.904 (3/4)	27.3659%	1.2865	<=1.417 (1/4)	4.0135%	0.4149
ESR_Onset	0.0219	2.3307%	0.2963	43.4187	<=44.6 (1/3)	58.8961%	1.1780	>68.5 (3/3)	20.4062%	0.6595
Dis_dur_at_MAS_onset	0.0218	2.3271%	0.2959	2.1457	>0.038 (2/2)	91.3226%	1.1021	<=0.038 (1/2)	8.6774%	0.5064
WBC_Onset	0.0211	2.2542%	0.2866	11.4340	<=1.626 (1/3)	9.5554%	2.0313	>5.927 (3/3)	59.7276%	0.9086

Tab 14. – MI of each variable for different outcome status

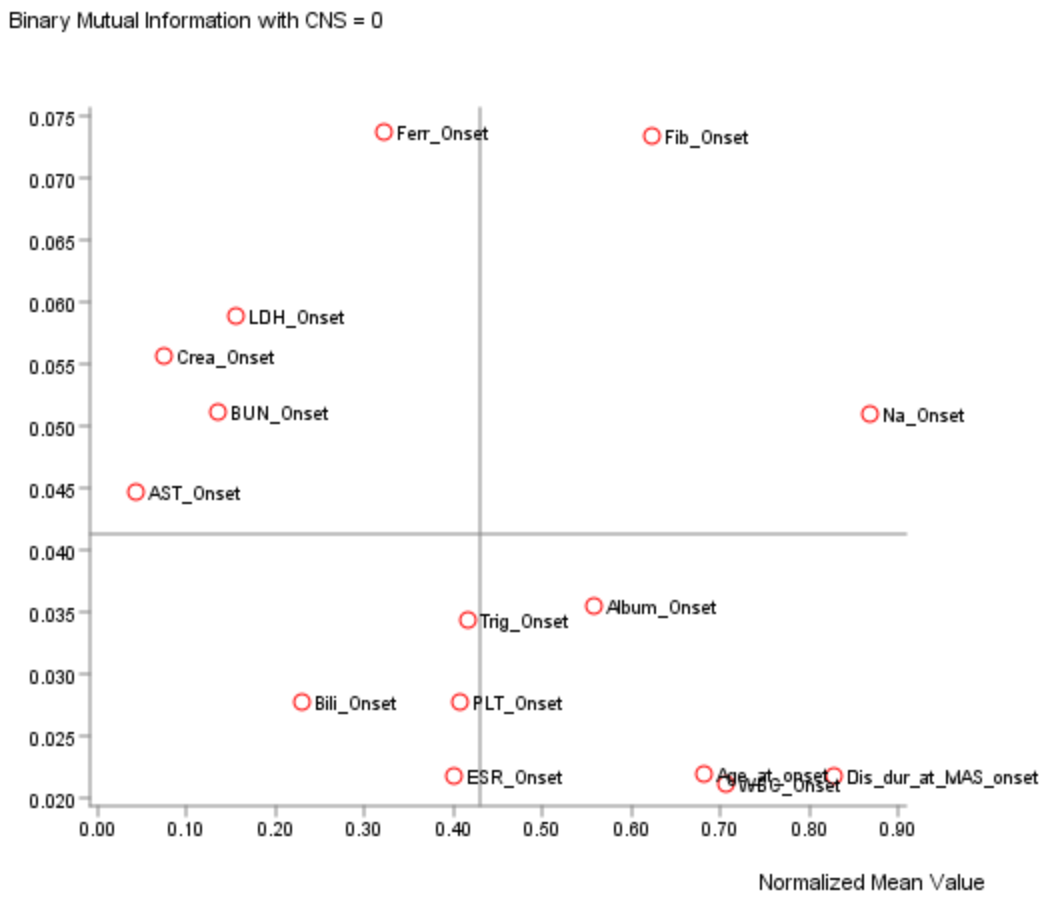


Fig 21. – Binary MI of each variable according to mean values

3.3.5 Target optimization: Which are the combinations of values associated with the highest probability of CNS disease?

Tab 15 shows the laboratory pattern at highest CNS-risk, characterized by hypofibrinogenemia and hypoalbuminemia. The target optimization tree and relative optimization score for CNS involvement are reported in Fig. 22.

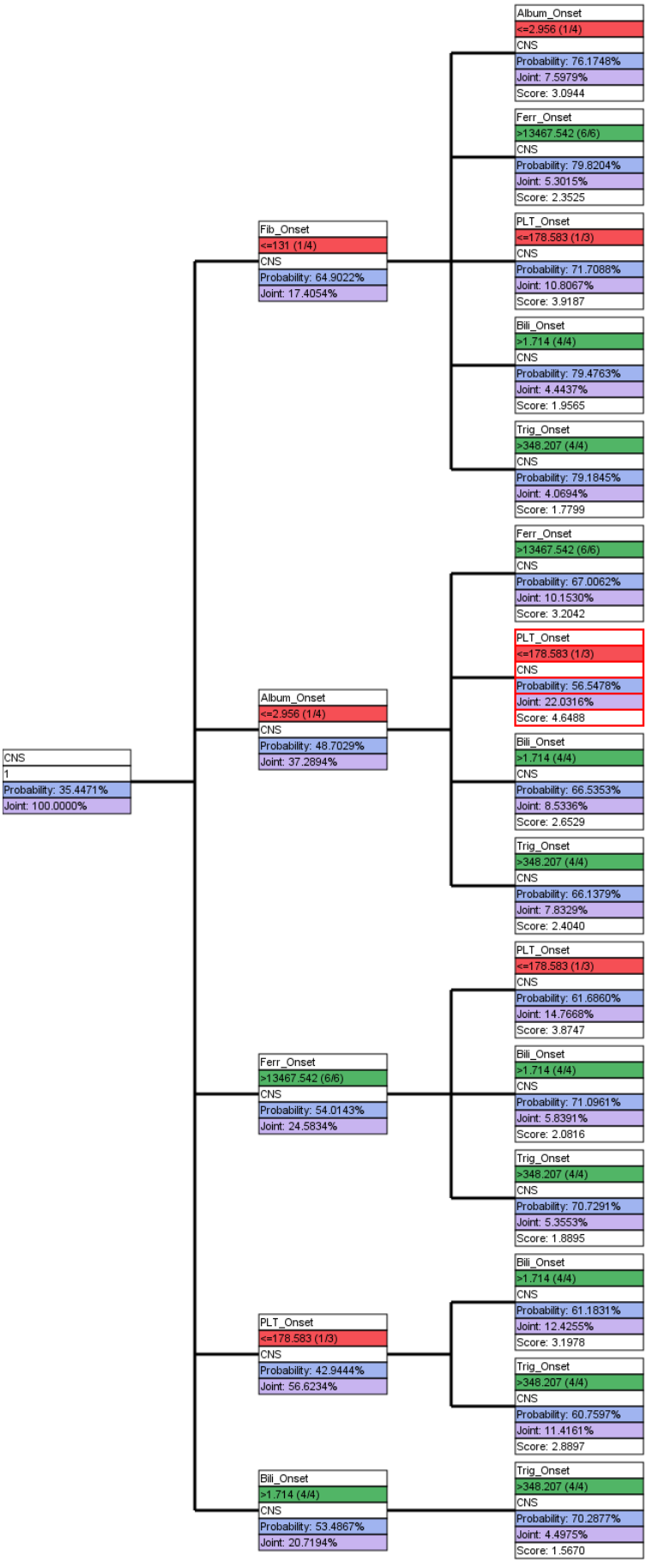
Dynamic Profile CNS: Probability Maximization (A posteriori)						
Search Method: Hard Evidence						
CNS = 0						
Node	Evidence	Posterior Probability P(s|E)	Marginal Likelihood P(E)	Likelihood P(E|s)	Bayes Factor BF(s,E)	Generalized BF GBF(s,E)
A priori		64.5529%	100.0000%			
Fib_Onset	>241 (4/4)	76.3802%	54.9641%	65.0345%	1.1832	1.5240
LDH_Onset	<=2118.227 (2/4)	83.5141%	34.8776%	45.1222%	1.2937	1.5352

Other Nodes		
Node	Prior Value/Mean	Posterior Value/Mean
Age_at_onset	6.6132	6.6014
Dis_dur_at_MAS_onset	1.9477	1.8896
WBC_Onset	12.3613	12.6336
PLT_Onset	214.4956	225.5881
ESR_Onset	51.1512	53.4225
AST_Onset	412.4229	323.6221
Trig_Onset	257.9873	246.9345
Na_Onset	135.2102	135.5425
Bili_Onset	1.4040	1.2783
Album_Onset	3.1603	3.2076
Ferr_Onset	12,111.1177	10,788.1627
BUN_Onset	30.9515	28.5013
Crea_Onset	0.6862	0.6270

CNS = 1						
Node	Evidence	Posterior Probability P(s|E)	Marginal Likelihood P(E)	Likelihood P(E|s)	Bayes Factor BF(s,E)	Generalized BF GBF(s,E)
A priori		35.4471%	100.0000%			
Fib_Onset	<=131 (1/4)	64.9022%	17.4054%	31.8686%	1.8310	2.2196
Album_Onset	<=2.956 (1/4)	76.1748%	7.5979%	16.3276%	2.1490	2.3732

Other Nodes		
Node	Prior Value/Mean	Posterior Value/Mean
Age_at_onset	6.6132	6.6385
Dis_dur_at_MAS_onset	1.9477	2.0726
WBC_Onset	12.3613	11.7762
PLT_Onset	214.4956	190.6695
ESR_Onset	51.1512	46.2726
AST_Onset	412.4229	603.1628
LDH_Onset	2,281.6461	2,803.7704
Trig_Onset	257.9873	281.7282
Na_Onset	135.2102	134.4966
Bili_Onset	1.4040	1.6741
Ferr_Onset	12,111.1177	14,952.7587
BUN_Onset	30.9515	36.2146

Search Method: Hard Evidence

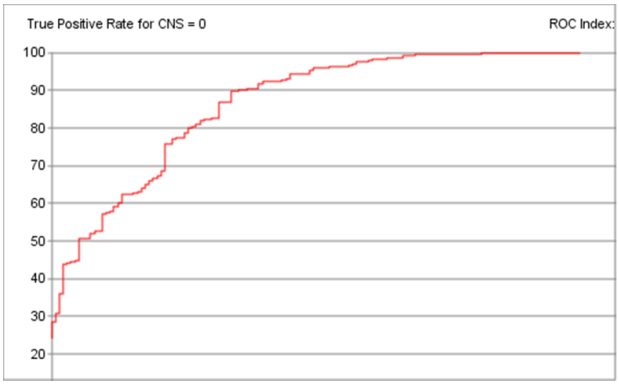


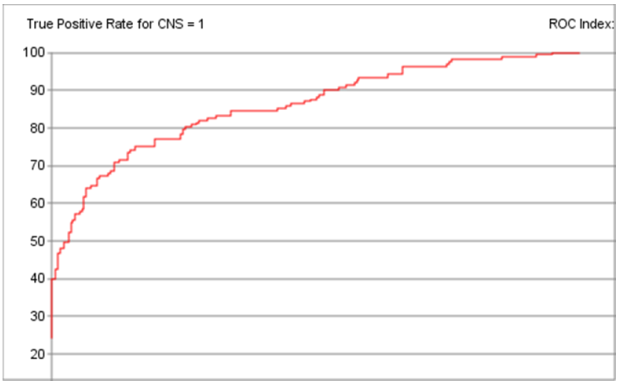
Node	State	Score
Album_Onset	<=2.956 (1/4)	14.4572
Fib_Onset	<=131 (1/4)	13.1020
PLT_Onset	<=178.583 (1/3)	12.3086
Ferr_Onset	>13467.542 (6/6)	10.6240
Bili_Onset	>1.714 (4/4)	6.5114
Trig_Onset	>348.207 (4/4)	5.2650

Fig. 22 – Target optimization tree and score for CNS disease prediction

3.3.6 Accuracy assessment

The model for CNS disease reached a c-index of 0.816, with a slightly worse curve for the absence of the outcome (Fig 23)





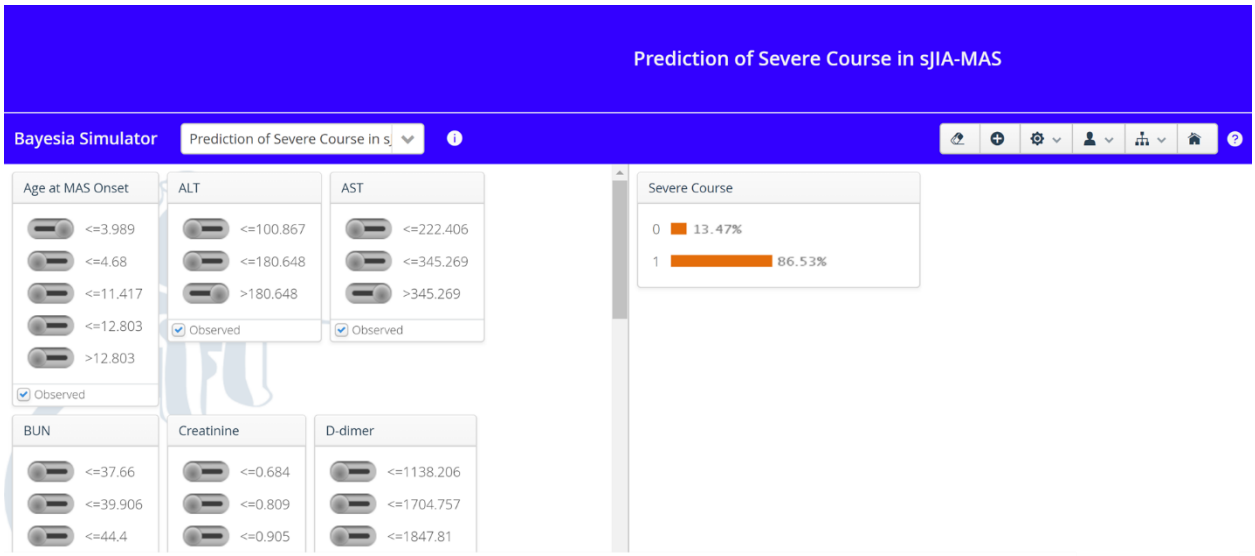
Target: CNS		
Value	0	1
Gini Index	23.5963%	41.3572%
Relative Gini Index	63.2730%	62.3805%
Lift Index	1.3143	1.7468
Relative Lift Index	89.6604%	83.7823%
ROC Index	81.6381%	81.1919%
Calibration Index	80.2942%	75.6368%
Binary Log-Loss	0.5467	0.5502

Overall Precision	78.7966%
Mean Precision	76.1176%
Overall Reliability	75.8289%
Mean Reliability	73.7099%
Overall Relative Gini Index	62.9610%
Mean Relative Gini Index	62.8267%
Overall Relative Lift Index	87.6056%
Mean Relative Lift Index	86.7213%
Overall ROC Index	81.4821%
Mean ROC Index	81.4150%
Overall Calibration Index	78.6661%
Mean Calibration Index	77.9655%
Overall Log-Loss	0.5031
Mean Binary Log-Loss	0.5484

Fig. 23 – ROC curves and accuracy indexes of the CNS disease model

4. Online prediction system for MAS outcome

Based on information theory metrics (conditional mutual information) from the BN model, a decision algorithm for the prediction of severe course was implemented and transposed in a web-based decision support tool for external users. Outcome probabilities predictions are



provided in real time according to the predictor values entered by the user. The prediction system is freely accessible at https://simulator.bayesialab.com/#!simulator/130366133559. Figure 24 shows a screenshot of the tool.

Fig. 24- Screenshots of an online prediction system for severe course in sJIA-MAS. The left-hand pane allows for entering predictor values for a given patient case, while the right-hand pane shows in real-time the associated probability of outcomes, according to the embedded TAN model.

5. Discussion

In this work we integrated our previous results to develop a prognostic model of sJIA-MAS in a probabilistic and causal inference framework, through the use of Bayesian Networks and information theoretic metrics. Bayesian network are powerful formalism to capture the complexity of the interrelationships between various biomarkers, representing dependency between causes and effects, and can account for the inherently uncertainty of the prognostic process in terms of probabilities, quantifying the risk in the form of probabilities[113]. Among the other advantages, BNs provide a visual and understandable way to represent the relationship between variables though an inference process that is more interpretable than most machine-learning based predictions, as it mimics human decision-making. Their potential in measuring risks under uncertainties has promoted their employment in different diagnostic and prognostic research[114] [115] [116].

Besides assessing the impact of single variables on the outcome, BNs can analyze how predictors are related to each other. This allows to simulate various scenarios and calculate the changes in outcome probabilities associated to a specific laboratory pattern, i.e. a combination of risk factors.

On the other hand, information theoretical measures allow to quantify the uncertainty in outcome (entropy) both prior to and following a laboratory test, providing a measure of the

[113] Lucas, Peter. Bayesian networks in medicine: a model-based approach to medical decision making. na, 2001.
[114] Langarizadeh, Mostafa, and Fateme Moghbeli. "Applying naive bayesian networks to disease prediction: a systematic review." *Acta Informatica Medica* 24.5 (2016): 364.
[115] Sesen MB , Nicholson AE, Banares-Alcantara R, Kadir T, Brady M. Bayesian networks for clinical decision support in lung cancer care. PLoS One. 2013;8(12):e82349.
[116] Cai ZQ , Si SB, Chen Cet al. Analysis of prognostic factors for survival after hepatectomy for hepatocellular carcinoma based on a Bayesian network. PLoS One. 2015;10(3):e0120805.

information gained by accumulating evidences[117]. For this reason, information theory has been employed in different clinical contexts and applied to the evaluation of laboratory data [118].

What emerged by BNs models and mutual information based algorithms is that hypofibrinogenemia , as well as hyponatremia and alteration of markers of kidney injury, namely creatinine and BUN, are the risk factors that have the highest impact on most severe outcomes and provide the greatest amount of information for the prediction of a complicated course and death. This is in line with previous evidences on the prognostic value of renal involvement in the HLH syndromes[119], and with results of our previous subgroups and causal analysis that employed regression modelling techniques.

A prognostic role, not previously highlighted, was demonstrated for LDH and triglycerides high levels, respectively for the forecasting of severe course and fatal outcomes, whose value as risk factors need to be investigated in further works.

Among other predictors, the current analysis confirmed the non-linear hazard conferred by ferritin values which, albeit not a causal driver of outcomes, becomes highly informative for the prediction of both severe course and CNS involvement above levels of 13000 ng/dl.

The model showed a good forecasting performance for all the outcomes. To allow for querying of the model in specific clinical scenarios, these results were embedded in an interactive web--based prediction system, avaible for external users, which is based on information theoretical

[117] Lee, Joon, and David M. Maslove. "Using information theory to identify redundancy in common laboratory tests in the intensive care unit." BMC medical informatics and decision making 15.1 (2015): 59.
[118] Vollmer R. Entropy and information content of laboratory test results. Am J Clin Pathol. 2007; 127:60–5
[119] Karras, Alexandre. "What nephrologists need to know about hemophagocytic syndrome." Nature Reviews Nephrology 5.6 (2009): 329.

measures. Traditional risk scores usually use coefficients from regression modelling to derive the risk weights for predictors. This, however, allows to account only for the individual relationship between the variable and the outcome in a linear way, ignoring the interrelationship among predictors. Other authors[120] [121]have developed Bayesian Network-based risk score for different conditions, generally showing an improvement in accuracy or usability compared with the original regression-based scores. In this study, we used the information encoded by the BN model on the effect on outcome for the key predictors and the information gain from each evidence, as calculated by the optimizing algorithm, to develop a BN-based simplified risk score. In this way, we aimed to incorporate in the score the probabilistic associations between variables. The score reached a decent c-index for severe outcome and showed a good accuracy for the prediction of mortality, outperforming the performance of the regression and CART modelling presented in Chapter 3.

Among the limitations of this study is the retrospective nature of the dataset. Second, the effects of treatments were not considered in this analysis. Moreover, we argue that for a comprehensive causal analysis of this syndrome, the inclusion of cytokine and immunologic mediators that are known to drive the pathogenesis of HLH/MAS is crucial. Although molecular and cytokine data were not avaible in this dataset, one of the advantages of the Bayesian Network formalism is that the model can be easily updated with further evidences, refined with expert knowledge and merged with evidences from independently performed but related studies. The integration

[120] Loghmanpour, Natasha A., et al. "A new Bayesian network-based risk stratification model for prediction of short-term and long-term LVAD mortality." ASAIO journal (American Society for Artificial Internal Organs: 1992) 61.3 (2015): 313.

[121] Kraisangka, Jidapa, Marek J. Druzdzel, and Raymond L. Benza. "A Risk Calculator for the Pulmonary Arterial Hypertension Based on a Bayesian Network." BMA@ UAI. 2016.

between routinely gathered laboratory parameters and cytokine levels could allow to estimate the probabilities of the latter in clinical context where they are unavailable, allowing for the identification, in probabilistic terms, of the most appropriate cytokinin target to address therapeutically, based on specific laboratory patterns. The potential demonstrated by this methodology encourages the validation and expansion of our analysis with a prospective and more complete data.

6. Conclusion

In summary, the combination of Bayesian networks and constructs from information theory allowed to identify the key determinants of sJIA-MAS outcomes. This study could represent a proof-of-concept showing the potential of this methodology in the context of complex disease like HLH/MAS. These results could help refine our approach to patient management, monitoring and stratification in MAS, to raise the awareness of mechanisms that drive the progression of MAS to unfavorable outcomes, and to rationally choose the most appropriate immunosuppressive and supporting therapies to improve the outcomes.

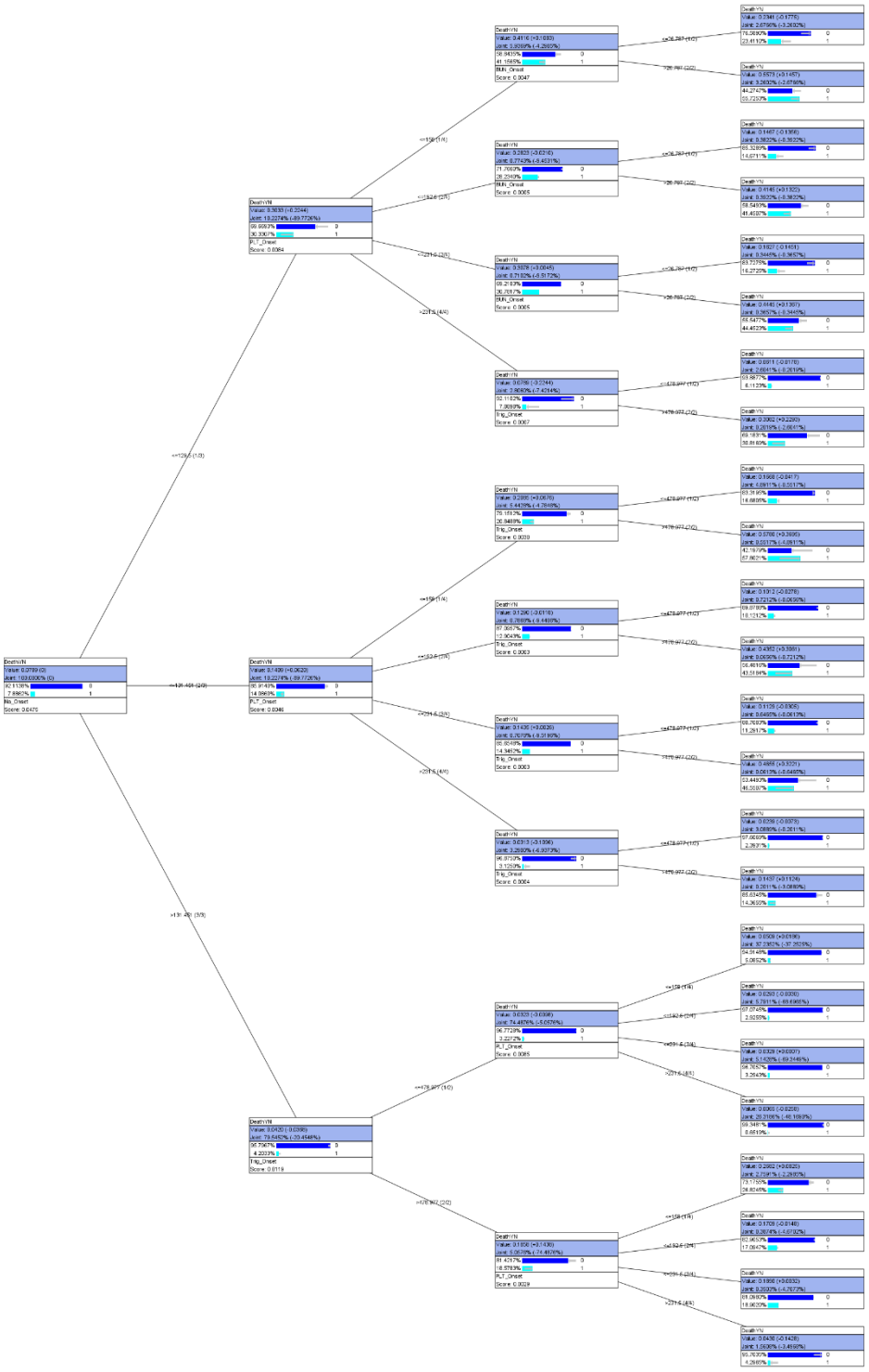
References

- Banu, Bazila, and Ponniah Thirumalaikolundusubramanian. "Comparison of Bayes Classifiers for Breast Cancer Classification." Asian Pacific journal of cancer prevention: APJCP 19.10 (2018): 2917.
- Ben-Gal I. Bayesian Networks. In: Ruggeri F, Faltin F, Kenett R, editors. Encyclopedia of Statistics in Quality & Reliability. 2013. Chichester: Wiley and Sons; 2007
- Cai ZQ , Si SB, Chen Cet al. Analysis of prognostic factors for survival after hepatectomy for hepatocellular carcinoma based on a Bayesian network. PLoS One. 2015;10(3):e0120805.
- Castillo, E., Gutierrez, J. M. & Hadi, A. S. (1997). Sensitivity analysis in discrete Bayesian networks,

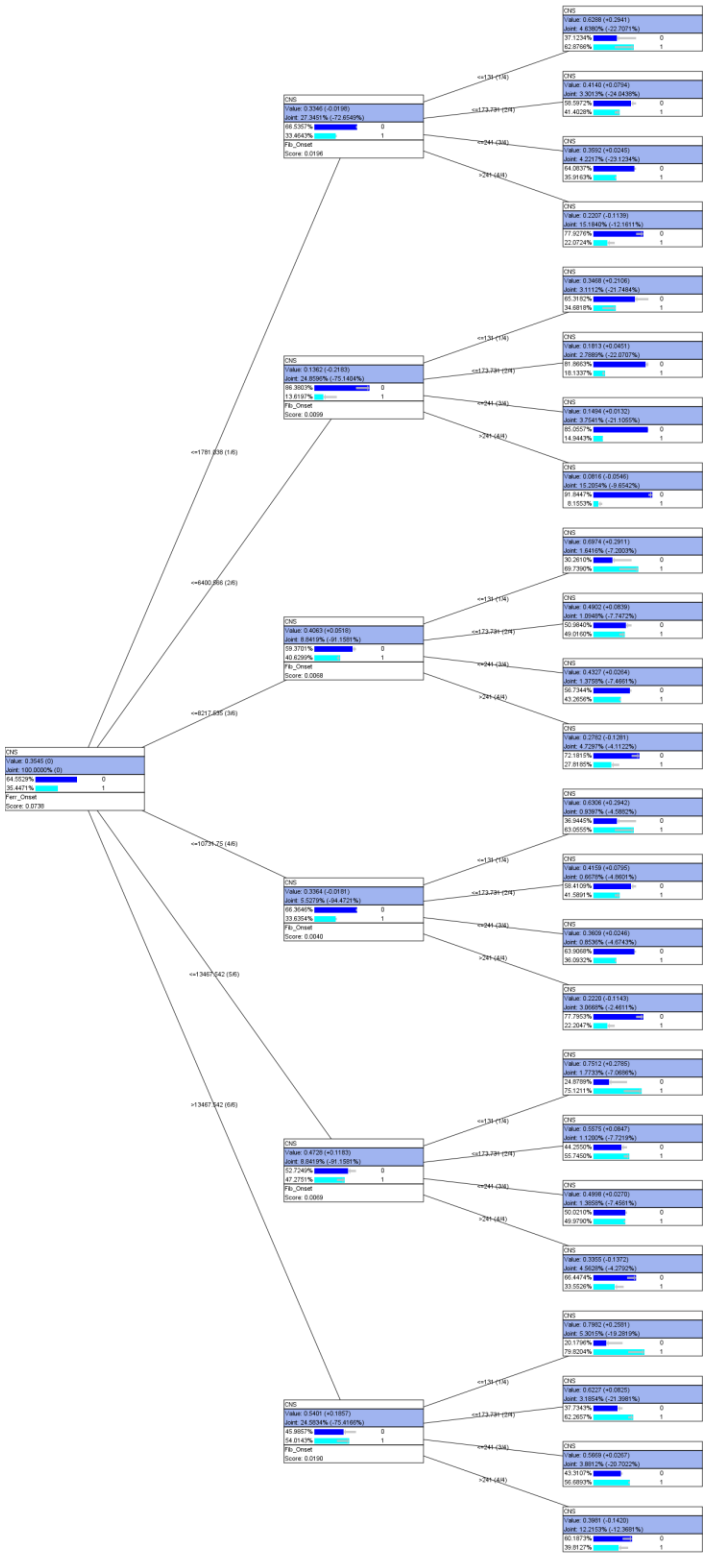
- Conrady, Stefan, and Lionel Jouffe. Bayesian networks and BayesiaLab: A practical introduction for researchers. Bayesia USA, 2015.
- Cox Jr, Louis Anthony, Douglas A. Popken, and Richard X. Sun. Causal Analytics for Applied Risk Analysis. Springer International Publishing, 2018.
- Delgado-Bonal, Alfonso, and Javier Martín-Torres. "Human vision is determined based on information theory." Scientific reports 6 (2016): 36038.
- Friedman N, Geiger D, Goldszmidt M (1997) Bayesian network classifiers. Machine Learning. IEEE Transactions on Systems, Man, and Cybernetics 27(4): 412-423.
- https://support.bayesfusion.com/docs/GeNIe/bn_sensitivitybn.html
- Karras, Alexandre. "What nephrologists need to know about hemophagocytic syndrome." Nature Reviews Nephrology 5.6 (2009): 329.
- Kraisangka, Jidapa, Marek J. Druzdzel, and Raymond L. Benza. "A Risk Calculator for the Pulmonary Arterial Hypertension Based on a Bayesian Network." BMA@ UAI. 2016.
- Kullback, S. (1959), Information Theory and Statistics, John Wiley & Sons. Republished by Dover Publications in 1968; reprinted in 1978: ISBN 0-8446-5625-9.
- Langarizadeh, Mostafa, and Fateme Moghbeli. "Applying naive bayesian networks to disease prediction: a systematic review." *Acta Informatica Medica* 24.5 (2016): 364.
- Learned-Miller, Erik G. "Entropy and mutual information." Department of Computer Science, University of Massachusetts, Amherst (2013).
- Learned-Miller, Erik G. "Entropy and mutual information." Department of Computer Science, University of Massachusetts, Amherst (2013).
- Lee, Joon, and David M. Maslove. "Using information theory to identify redundancy in common laboratory tests in the intensive care unit." BMC medical informatics and decision making 15.1 (2015): 59.
- Lee, Joon, and David M. Maslove. "Using information theory to identify redundancy in common laboratory tests in the intensive care unit." BMC medical informatics and decision making 15.1 (2015): 59.
- Loghmanpour, Natasha A., et al. "A new Bayesian network-based risk stratification model for prediction of short-term and long-term LVAD mortality." ASAIO journal (American Society for Artificial Internal Organs: 1992) 61.3 (2015): 313.
- Lucas, Peter. Bayesian networks in medicine: a model-based approach to medical decision making. na, 2001.
  markers for breast cancer. In Proceedings of the 2011 IEEE International Conference on Industrial Engineering and Engineering Management, Singapore. Piscataway, IEEE: 1826–1830
- Minoia, Francesca, et al. "Clinical features, treatment, and outcome of macrophage activation syndrome complicating systemic juvenile idiopathic arthritis: a multinational, multicenter study of 362 patients." Arthritis & Rheumatology 66.11 (2014): 3160-3169.
- Minoia, Francesca, et al. "Dissecting the heterogeneity of macrophage activation syndrome complicating systemic juvenile idiopathic arthritis." The Journal of rheumatology 42.6 (2015): 994-1001.

- Minoia, Francesca, et al. "Dissecting the heterogeneity of macrophage activation syndrome complicating systemic juvenile idiopathic arthritis." The Journal of rheumatology 42.6 (2015): 994-1001.
- Nicholson, Ann E., and Nathalie Jitnah. "Using mutual information to determine relevance in Bayesian networks." Pacific rim international conference on artificial intelligence. Springer, Berlin, Heidelberg, 1998.
- Ravelli, A., et al. "Macrophage Activation Syndrome." Handbook of Systemic Autoimmune Diseases. Vol. 11. Elsevier, 2016. 85-106.
- Sesen MB , Nicholson AE, Banares-Alcantara R, Kadir T, Brady M. Bayesian networks for clinical decision support in lung cancer care. PLoS One. 2013;8(12):e82349.
- Si S, Liu G, Cai Z, Xia P (2011) Using Bayesian networks and importance measures to identify tumor
- Thornley S, Marshall RJ, Wells S, Jackson R (2013) Using Directed Acyclic Graphs for Investigating Causal Paths for Cardiovascular Disease. J Biomet Biostat 4: 182. doi:10.4172/2155-6180.1000182
- Tsagaris, Vassilis, Nikos Fragoulis, and Christos Theoharatos. "Performance evaluation of image fusion methods." Image Fusion (2011): 71-88.
- Vollmer R. Entropy and information content of laboratory test results. Am J Clin Pathol. 2007; 127:60–5
- Wikipedia contributors. (2019, December 13). Kullback–Leibler divergence. In Wikipedia, The Free Encyclopedia. Retrieved 14:33, January 8, 2020, from https://en.wikipedia.org/w/index.php?title=Kullback%E2%80%93Leibler_divergence&oldid=930573126

Appendix



-Target interpretation tree for the prediction of mortality based on mutual information



-Target interpretation tree for the prediction of CNS disease based on mutual information

Conclusions

Macrophage activation syndrome is a fearsome complication of many autoimmune conditions, particular associated to Systemic Juvenile Idiopathic Arthritis. The clinical and laboratory spectrum of its manifestation is wide, encompassing dysfunction of different organs, and the highly heterogenous in its severity. The course of the disease can be unpredictable, in some case rapidly evolving in multi organ failure and fatal outcome. Recent research efforts have focused on improving the accuracy and promptness of diagnosis, through the development of diagnostic criteria. However, despite the recognized severity, prognostic research on this condition is still at an early stage. To get insights on the dynamics of disease progression and discover which risk factors can help predict poor outcomes at disease onset, we combined subgroup analysis, regression modelling and Bayesian methodology within a causal inference and probabilistic framework for the analysis of a large multinational cohort. We identified selected laboratory alterations at MAS onset, among the test that are routinely obtained in these patients, which are surrogate for the most probable determinants of disease progression toward unfavorable course and are most informative for outcome prediction. Based on these findings, we developed a probabilistic prognostic model for sJIA-related MAS that is able to forecast with a good accuracy the clinical evolution of these patients and provides personalized predictions through a web-based predictive tool. Further research is warranted for the validation of these results and the expansion of the model with further predictors that can improve its performance. These findings add to the current understanding of MAS and, through a better stratification of patients and identification of appropriate targets of clinical monitoring and treatment, can help improve the outcomes of this complex and severe condition.

Acknowledgements

I would like to thank my supervisor, Prof. Angelo Ravelli, for the opportunity to conduct this research and PhD programme. I would like to express my gratitude also to Dr. Nicolino Ruperto for supporting my interest in methodology. A special thank goes to Dr. Alessandro Consolaro for his fundamental support and guidance during my doctoral years. Thank you to all of my colleagues and PRINTO research assistants.

I would also like to remember here Aaron Swartz and Alexandra Elbakyan, whose works made this research, as well as many others worldwide, possible.

Lastly, I would like to thank my parents, my sister Alessia and Luke.